Expert Systems With Applications 192 (2022) 116300

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Review

# Data association in multiple object tracking: A survey of recent techniques

Lionel Rakai, Huansheng Song, ShiJie Sun *, Wentao Zhang, Yanni Yang

*Department of Intelligent Transport and Information Engineering, Chang An University, Middle Section of Nanerhuan Road, Xi'an City, Shaanxi Province, 710064, China*

ARTICLE INFO

*Keywords:*
Data association
Probabilistic association techniques
Hierarchical association techniques
Interactive multiple model
Kalman filter
Multiple object tracking

ABSTRACT

The advances of Visual object tracking tasks in computer vision have enabled a growing value in its application to video surveillance, particularly in a traffic scenario. In recent years, significant attention has been made for the improvement of multiple object tracking frameworks to be effective in real-time while maintaining accuracy and generality. By breaking down the tasks involved in a Multiple Object Tracking framework based on the Tracking-By-Detection approach — an extension of simply detecting and identifying objects, further involved solving a filtering problem by defining a similarity function to associate objects. Hence, this paper focuses on the task of data association via uniquely defined similarity functions and filters only where we review current literature about these techniques which have been used to advance the performance in MOT for vehicle and pedestrian scenarios. While there is difficulty in classifying the quantitative results for the association task only within a proposed MOT framework, our study tries to outline the fundamental ideas put forward by researchers and compare results in a theoretically qualitative approach. Tracking methods are reviewed by categories based on legacy techniques like Probabilistic and Hierarchical methods, followed by an analysis of new approaches and hybrid models. The models identified in each category are further analysed based on performance in stability, accuracy, robustness, speed and computational complexity to derive an understanding of which direction the research within the data association level is strong and which is lacking. Our review further aims to identify the successful models applied to recognize the weaknesses for future improvement.

## 1. Introduction

Multiple object tracking (MOT) has become an important area in video sequences for traffic surveillance (Kokul et al., 2015; Lee et al., 2017a; Zeng et al., 2016), security monitoring (Gong, 2005), behaviour analysis (Dehghan et al., 2015; Dimitrievski et al., 2019), action recognition (Cherian et al., 2018; Choutas et al., 2018; Luvizon et al., 2018), etc. Including the task of associating individual objects, initiating, maintaining and correctly terminating the track of an object has become paramount for the overall improvement in the performance of an MOT framework (Bergmann et al., 2019; Milan et al., 2013; Noh et al., 2015; Sun et al., 2019). Taking previously known information about an objects shape (Lee & Hwang, 2015; Steyer et al., 2018), movement (Anuj & Krishna, 2017; Lee & Hwang, 2015; Lin & Hung, 2018), poses (Dorai et al., 2017; Tang et al., 2018a; Tang & Hwang, 2019) and changes in appearance (Rasmussen & Hager, 2001; Taalimi & Qi, 2015; Tran & Harada, 2013), the process of data association involves comparing this previously learned information about newly identified objects within an input video frame. While data association can cover information about identification, location and trajectory, this study focuses on associating objects trajectory-based tasks. Issues often associated with data association include missing detections, occlusion and target interaction within crowded scenes.

Using the information that was previously observed based on the track, pose or identity to match with new information about objects, algorithms have been developed for optimization and mathematically add value for computer vision researchers. The current weaknesses are found in trying to improve the accuracy and efficiency of introducing and recovering tracks without increasing the rate of false alarms.

Optimizing a data association algorithm, filter or function involves identifying key challenges in the physical condition of the input frames such as different levels of noise (Tang et al., 2018b; Wang et al., 2018b; Yang et al., 2018b), clarity (Zhu et al., 2017) and low to medium image resolutions (Gao et al., 2015; Jiang et al., 2015; Tang et al., 2018b). Handling the spatial and temporal state conditions (Seong & Park, 2012) is another challenge which includes the popularly mentioned partial and full occlusion states, frequent occlusion occurrences, illumination, object re-identification over short and longer periods of

* Corresponding author.
*E-mail addresses:* lionelrakai@ymail.com (L. Rakai), hshsong@chd.edu.cn (H. Song), shijieSun@chd.edu.cn (S. Sun), 2018024001@chd.edu.cn (W. Zhang), yyanni@163.com (Y. Yang).

https://doi.org/10.1016/j.eswa.2021.116300
Received 19 November 2019; Received in revised form 24 November 2021; Accepted 25 November 2021
Available online 18 December 2021
0957-4174/

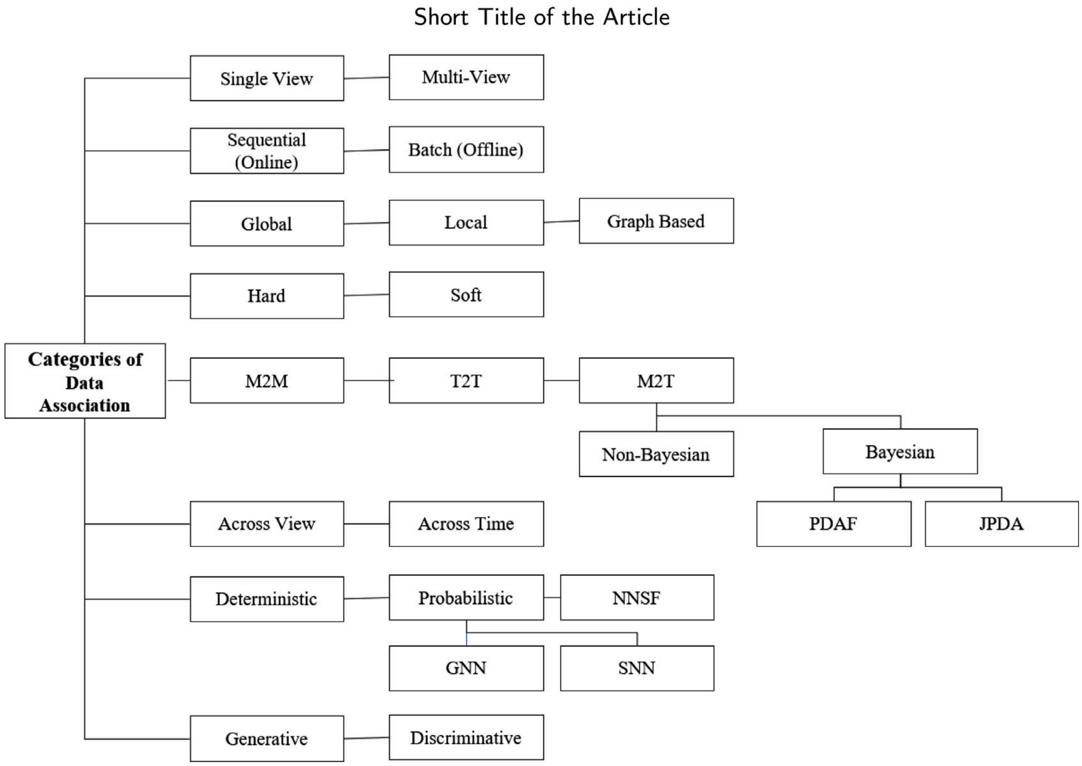


**Fig. 1.** A classification summary of the data association tasks used in previous work, including mentions in the Related Works sections of published papers. Most classifications are divided into two areas, for example, Local and Global, Single View and Multi-View, Hard and Soft Associations, to name a few. Since classifications have been well described in Object Tracking literature, our paper focuses on the new techniques applied in popular classifications.

time, scale changes, motion blur due to fast movement and adaptation to rotational appearance changes. While these challenges are often described for improvement in multiple object tracking tasks, it is specifically the layer of data association that needs to handle these encounters for optimization. In reference to considering the spatial and temporal states, Seong et al. divided the association task into four aspects — distinguishability, visual feature, spatiotemporal prediction and change prediction to illustrate a proposed likelihood model and the intermediate relationships.

Though a future concept may involve the merging of tasks for detection and association within an MOT framework, at present, it is observed that the task of data association is still predominantly separated into a specified model or layer. The scope of this paper applies to the survey of work published within the last three years and filtered by area of interest. The areas include computer vision, computer science and visual multi-object tracking for vehicles and pedestrians in a traffic surveillance video, publication year and keywords. Considerations were also made in terms on the number of citations made per paper and the journal or source of publication. The individual related works categorize association methods in many forms which have been summarized in Fig. 1, however, this review focuses on the type of technique rather than the category because of the increasing effort to merge advantages of multiple categories in order to offer a more generalized model application. For example, sequential (Al-Shakarji et al., 2018; Hou, Yang et al., 2017; Ritter et al., 2018) and batch (Taalimi & Qi, 2015; Yang et al., 2018a), local and global (Bozorgtabar & Goecke, 2017; Steyer et al., 2018; Xiao & Zhong, 2017), hard (Godinez & Rohr, 2014; Niedfeldt et al., 2017) and soft (Chen et al., 2017a; Date et al., 2014) approaches.

Due to the multiplied number of possible technique types, models are categorized and discussed based on its legacy techniques — Probabilistic (Chalvatzaki et al., 2018a; Huang et al., 2017; Rasmussen & Hager, 2001), Hierarchical (Chen et al., 2018; Li et al., 2018; Liu et al., 2017), IMM (Blackman et al., 1995; Yuan et al., 2017b), Kalman Filter Based (Jeong et al., 2014; Milan et al., 2013; Sahbani & Adiprawita, 2016), Fuzzy Association (Li et al., 2017a; Tafti & Sadati, 2010; Xiyang et al., 2018) and New Techniques in particular. Since a numerical qualitative results comparison cannot be performed without considering the elements within discussed frameworks, we instead highlight key feature performance characteristics which are met by each model. These characteristics include stability, accuracy, speed, robustness and computational complexity.

The remainder of this paper is arranged in the following order: Section 2 describes the previously surveyed work and defines the categories of data association in the existing literature. Section 3 presents the recent techniques of data association applied in current Multiple Object Tracking frameworks for pedestrian and vehicle environments. Section 4 offers a qualitative summary table and a discussion of the updated techniques. Section 5 contains the conclusions.

## 2. Previously surveyed work

Since the function of data association is embedded as an element within the Multiple Object Tracking frameworks, there are more comparative reviews on full tracking frameworks (Fan et al., 2016; Fiaz et al., 2018; Li et al., 2018; Luo et al., 2014; Mandal & Adu-Gyamfi, 2020; Ooi et al., 2020; Wang et al., 2018b) rather than the individual elements. Recent reviews include a survey on tracking algorithms by Fan et al. (2016), a review of tracking methods for noisy targets by Fiaz et al. (2018) and an MOT literature review by Luo et al. (2014) who covers the processes, components, models and evaluations for frameworks designed to track multiple objects. Fiaz et al. (2018) lists and describes specifically published filters and trackers that try to overcome different levels of noisy data while Fan et al. (2016) offers a simple understanding of the Visual MOT process and existing categorizations. Recent reviews based on data association include a different approach on classification with three traditional data association methods (probability, conversation and layering) and then focuses on deep learning based methods by Li et al. (2020). Yarkony et al. (2020) surveys data association methods where an instance is based on parameters by a set of observations and a set of possible hypotheses.

Some specific analysis on data association techniques have been documented, particularly by Buluttekin et al. (2017) in the area of multiple vehicle tracking but limits its study to the identification of the GNN and PDAF algorithms for comparison. Our study, in particular, tries to cover as many classifications of association algorithms published in recent years as well as considering methods that apply to both pedestrians and vehicles alike. Rasmussen and Hager (2001) surveyed methods only under the Probabilistic Data Association umbrella and its evolution to Joint approaches such as the JPDAF (Joint PDAF) and JLF (Joint Likelihood Filter). Hou, Wan et al. (2017) published a review of pedestrian tracking by comparing methods between single and overlapping cameras. Another notable survey performed under the subject of object tracking includes a review of RGB-D video datasets by Zhang et al. (2016). Emami et al. (2020) compares measurement-to-track and track-to-track type of associations specifically.

Unique categories of data association that have been defined by past works are listed below:

1. Single View and Multi-View: Single view refers to associations related to data obtained from a single camera or sensor angle in a specific time-frame while multi-view refers to a single scene and a single timestamp with multiple camera angle view options and multiple sensors. The multi-view approach does offer an advantage when training an appearance model for objects at risk of running with low computational efficiency.
2. Sequential (online) and Batch (offline): these are classification methods under the tracking-by-detection approach where batch tracking methods build multiple tracks by analysing and optimizing the entire video or an entire sliding window in an offline mode. methods have been better applied to real-time applications because tracks are built on a frame-by-frame basis, and data association can be performed between successive two frames at a time.
3. Global, Local and Graph-based: Global data association is similar to the batch method in the sense that processing is performed over a batch of frames rather than individual sequential frames also making it difficult to apply to online streaming and processing. Local association applies to a smaller environment with less room for generalization as compared to global methods. In the case of graph-based methods, the track graph-based approach was introduced into multiple object tracking due to its effectiveness in long term tracking of objects and features. Chong (2012) wrote a review of graph approaches in data association specifically.
4. Hard and Soft: Hard and soft data association relates to the source of data being either hard or soft. Hard data is connected to quantitative features, while soft refers to qualitative. These terms also refer to the type of decisions made; for example, PDAF uses soft decisions by averaging out all the association possibilities. NNSF (Nearest Neighbour Standard Filter) is an example that takes hard decisions in a greedy approach.
5. M2M, M2T and T2T: Michaelis et al. (2017) explains that M2 m (Measurement to Measurement Association) uses all available information from individual measurements for an association by relating measurements from one sensor to measurements originating from another sensor. M2T (Measurement to Track Association) the data from one sensor at a single timestamp is associated with the tracked object, whereas for T2T (Track to Track Association) each sensor tracks the object individually.
6. Bayesian and Non-Bayesian: we would expect the analysis or algorithm that simply applies the Bayesian Theorem to be termed a Bayesian approach, but it is not completely similar to Bayesian statistics. In Artificial Intelligence, Bayesian-based networks is also another name for graphical-based models or Belief Networks (BN) where Bayes' rule is used for probabilistic inference. It is therefore applied in applications of Probabilistic Data Association (PDA). Non-Bayesian approaches are based on hypothesis testing and connected with the frequency of features or events to confirm a particular model. Based on the field of data association and MOT, Non-Bayesian methods are likelihood-based where interpretation tends to be more subjective.
7. NNFS, SNN and GNN: Nearest Neighbour and Finite Sets are common terms used under Probabilistic Data Association for target tracking due to their strength with sparse representations and a combination with Finite Sets such as PHD (Probability Hypothesis Density) filters try to estimate the target state without association. While most association methods use a measure of probability to evaluate different hypotheses, the Global Nearest Neighbour attempts to find the single most likely hypothesis within each scan of a frame-by-frame or batch-by-batch analysis. The SNN (Suboptimal Nearest Neighbour) assignment algorithm uses a probabilistic method which assigns observations to existing tracks and minimizing some distance criteria.
8. PDAF and JPDA: PDAF (Probabilistic Data Association Filter) uses a weighted average of all the measurements within a particular tracks validation region and is more common with single target tracking. The JPDA or Joint approach is an extension of the PDAF which applies to multi-target tracking.
9. Across-View and Across-Time: Across-view data association matches objects which appear in different views of multiple cameras while across-time data association matches currently observed objects with previously observed object tracks.
10. Deterministic and Probabilistic: Deterministic data association defines the optimal track set by global optimization algorithms and work with prior knowledge of object characteristics, making it quite useful in vehicle detection. Probabilistic methods apply the Minimum Mean Square Error (MMSE) Estimate to confirm or terminate tracks.
11. Generative and Discriminative: These refer to different approaches of matching, and they tend to refer to classifications under Bayesian statistics. Generative methods have gained popularity with their ability to exploit unlabelled data in addition to labelled data. However, this case is removed if a model needs to be trained discriminatively to improve generalization. Discriminative approaches offer better generalization but require completely labelled data.

## 3. Techniques

The following section groups more specific approaches for techniques that have become popular solutions within tracking frameworks in recent years. A break down of techniques used are complemented with supporting works that have applied these solutions (see Fig. 2).

### 3.1. Taxonomy

The classification of techniques has been categorized based on the association structure and mechanism, algorithms used and its underlying theoretical structure. Previously, the condensation of taxonomies for data association methods has been applied based on filter types such as Kalman and Bayesian, or association of probabilities alone (PDA, JPDA). However, when considering the algorithms that do not apply to probabilistic methods such as LSTM, it generalizes to different conforms such as hierarchical methods. Hierarchical and probabilistic methods, in particular, have not been compared side by side in previous works. In contrast with these two legacy methods, new and further updated/modified techniques are separately identified and discussed to illustrate an efficiency comparison in the field of Multiple Object Tracking and to also verify if the popularity of an algorithm is directly proportional to its effectiveness, particularly in the area of intelligent transportation.

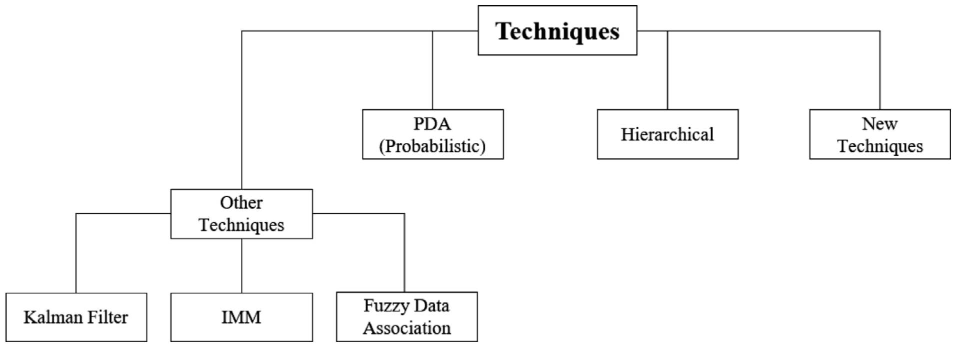


**Fig. 2.** The categories described in this paper have been simplified for the cleaner organization of papers and have been grouped among the different classifications illustrated. Apart from Probabilistic and Hierarchical methods, other categories have been further subdivided while new techniques are grouped together.

**Table 1**
Results for 2-D Real Image Sequences In Terms of Tracking Accuracy demonstrating greater accuracy by Kalman Filter but more consistency among the sequences for the Particle Filter (Godinez & Rohr, 2014).

| Sequence | Particle filter | Kalman filter |
|---|---|---|
| 1 | 73 | 96 |
| 2 | 55 | 81 |
| 3 | 63 | 90 |
| 4 | 67 | 82 |
| 5 | 94 | 94 |
| 6 | 64 | 68 |
| 7 | 82 | 86 |
| 8 | 74 | 77 |
| 9 | 70 | 81 |
| Mean | **71** | **84** |
| Std. Dev. | **11** | **9** |

### 3.2. Probabilistic

The Probabilistic aspect refers to the Bayesian information (Xiang et al., 2019; Yoon et al., 2015) while the filtering tracking algorithm assumes the state and measurement equations to be linear. A Bayesian strategy applied with Unmanned Aerial Vehicles (UAV) (Barkley & Paley, 2017) implemented but requires more optimization with improved motion models. In this respect, methods can be distributed between the Kalman Filter (Jeong et al., 2014; Li et al., 2018; Sahbani & Adiprawita, 2016) and Particle Filter (Chalvatzaki et al., 2018b; Kokul et al., 2015; Yingyi et al., 2017). The Kalman Filter considers single measurements per object based on a bottom-up localization algorithm, applied recently in Luo et al. (2019) while the particle filter can query multiple image positions to determine the location of an object at the expense of high computational cost. Godinez and Rohr (2014) published a tracking accuracy results table comparing the particle and Kalman filter algorithm on microscopy images. Table 1 (Godinez & Rohr, 2014) displays multiple accuracy results for 2-D Real-Time Image Sequences in terms of tracking accuracy, but we have displayed only a comparison between the Particle Filter and Kalman Filter. The results show a higher mean and standard deviation for the Kalman Filter, however, suggestions do state how particle filter is more useful in situations of tracking multiple objects.

Zhong et al. (2016) uses the particle filter data association developed into a multi-mode method to approximate target posterior distributions for non-linear systems to improve detection and tracking accuracy. A modification of particle filters is the Probability Hypothesis Density (PHD) filter where further development was made by Leonard and Zoubir (2019) into a Diffusion Particle and Multi-Sensor versions where the first was tested to track new targets faster correctly. Chalvatzaki et al. (2018b) merged particle filters with PDA and an Interactive multiple models for real-time selection of accurate motion models and providing more robust estimates.

Variational Bayesian (VB) (Lan et al., 2016; Tang & Hwang, 2019; Zhu et al., 2017) was formulated based on analytical approximations to posterior distributions and by modelling the tracks with a state space approach it offers the advantage of linear computational complexity and confirms better results over the existing commonly used MHT algorithm (Xiang et al., 2019). Further development by Lan et al. (2016) for solving joint multi-mode detections is called the Joint Detection and Tracking (JDT-VB) method. If moving to a non-linear state, the extended version of the Kalman Filter explained in (Ribeiro, 2004) is applied as the base algorithm (EKF) (Jiang & Cao, 2016; Mei et al., 2017; Yuan et al., 2017a). The Probabilistic Data Association Filter (PDAF) (Rasmussen & Hager, 2001; Ritter et al., 2018) was derived from the Kalman filter to overcome weaknesses when tracking single objects by introducing a notion of combined innovation. However, when considering the tracking of multiple objects, it is not as convenient to run a single PDAF tracker for each object (Rasmussen & Hager, 2001). Branching from the use of Bayesian information include the Permutation Matrix Track Association (PMTA) (Lee, Kanzawa et al., 2018), Global Nearest Neighbour (GNN) (Haag et al., 2018; Steyer et al., 2018), Particle Filters (Kara & Özkan, 2018; Piao et al., 2016), Multiple Hypothesis Tracking (MHT) (Chen et al., 2017b; Hou, Yang et al., 2017; Sheng et al., 2018; Yoon et al., 2018), Deep Person Re-Identification (Babaee et al., 2018; Guo & Cheung, 2018; Meng et al., 2019; Shen et al., 2018) and the still popular Probability Data Association Filter (PDAF).

According to Bar-Shalom et al. (2009) there are eight assumptions to summarize the PDAF algorithm — only one target of interest is present, the track has been initialized, the past time information about the target is summarized in the form of the Gaussian posterior (Kaiser et al., 2018; Lázaro-Gredilla et al., 2012), a measurement validation region is set up around the predicted measurement at each time, if the target is detected, and the measurement falls in the validation region than at most one of the validated measurements can be target originated, the remaining measurements are assumed to be false alarms, and the target detections occur independently over time with a known probability.

Extensions under the branch of PDAF include the PDAE elliptical approach (Godinez & Rohr, 2014) which interprets the association probabilities of each measurement as weights relative to the image likelihood of the object and Joint Integrated Probabilistic Data Association Filter (JIPDAF) (Hunde & Ayalew, 2018). A documented enhancement of the PDAF is the Joint-PDAF which allows for more advantages in multiple target tracking by introducing an exclusion principle that avoids the scenario of two or more trackers connecting to the same target. Further extensions to JPDAF have been proposed, including Smooth Variable Structure Filter (SVSF-JPDA) (Luo et al., 2019), Markov Chain JIPDA (MCJIPDA) (Lee et al., 2017b), Multi-Frame JPDA (MFJPDA) (Hamid Rezatofighi et al., 2015), JPDA based on extended target tracking (ETT) (Vivone et al., 2015) and Multi Hypothesis JPDA (MHJPDA) (Stauch et al., 2017). Rasmussen and Hager (2001) highlighted the disadvantages of JPDAF, such as the


**Fig. 3.** In a study of Probabilistic Data Association Methods documented by Rasmussen and Hager (2001), there was an illustration of tracking results across homogeneous regions for PDAF and JPDAF methods. Clearly, the PDAF model fails to maintain the track once the object has been occluded for a short period of time, indicating an issue with re-identification.

requirement for every tracker to have the same image likelihood and also the measurement process generates problems when objects overlap. The paper also illustrates the results difference between PDAF and JPDAF when tracking across homogeneous regions (Fig. 3) (Rasmussen & Hager, 2001). A sample-based approach was proposed by Zhang and Tang (2018) to improve state estimation and for accuracy improvement in the presence of inter-object occlusion. The Joint Likelihood Filter (JLF) was developed to encode the measurement association and the likelihood of the measurement for all targets. He et al. (2021) applies the JPDA with a graph matching method to meet the demands of long term tracklet association while maintaining accuracy in crowded scenes though it still adds on to computational complexity. This computational cost was addressed by Qu et al. (2020) by applying JPDA as a reinforcement learning based association approach.

Cabrera (2018) applied the Global Nearest Neighbour association technique to schedule variable field-of-view sensors for tracking multiple objects while multiple pedestrian tracking has been exploring Deep Person Re-Identification (Meng et al., 2019; Wang et al., 2018b) while Zhang et al. (2016) introduced a deep self-paced learning algorithm. Re-ID features have also been incorporated into likelihood association models for multiple pedestrian tracking to improve the tracking performance qualitatively.

Altendorfer and Wirkert (2016) analysed the Measurement-to-track association method by comparing and outlining the limitations of the Mahalanobis algorithm, such as the number of uncertainties. These are rectified by deriving an association log-likelihood distance which is used in assignment algorithms, and this works better in steady-state scenarios. To consider the effects of camera motion such as translation, pitch or raw motion of the camera, Yoon et al. (2019) proposed a framework of data association in two steps which inferred a structural constraint event aggregation method followed by recovery of missing objects between frames. Structural constraints are utilized by introducing a new cost function.

If the number of tracking targets varies, such as with human targets, the Probability Hypothesis Density filter (PHD) has the advantage of estimating the cardinality of targets and their states. To recognize the interaction and typical behaviour of human targets, a social force model (SFM) (Feng et al., 2016) was incorporated within the MCMC chain and paired with the PHD filter which improved localization and cardinality.

Based on general analysis, there has been more focus on accuracy in recent years (Cabrera, 2018; Chalvatzaki et al., 2018b; Granström, Svensson et al., 2017; Hunde & Ayalew, 2018; Lan et al., 2016; Lee et al., 2017b; Leonard & Zoubir, 2019; Luo et al., 2019; Mahemuti et al., 2016; Rasmussen & Hager, 2001; Ritter et al., 2018; Stauch et al., 2017; Yoon et al., 2019; Zhang & Tang, 2018) over the real-time requirement of speed and computational complexity. Since 2020, some research has demonstrated improved computational cost alongside good accuracy performance such as Doherty et al. (2020), He et al. (2020b), Qu et al. (2020) and Rangesh et al. (2021) While improvements in robustness have been made under the Probabilistic approach, there is still more attention required to generalizing association models and reducing the reliability on better defined motion models. Weaker areas of interest that are rarely highlighted include cardinality and overall stability. Fig. 4 illustrates the breakdown of taxonomy (see Table 2).

### 3.3. Hierarchical

Hierarchical association methods take advantage of layering to allow multiple techniques for different tasks. Recent approaches have been spread between the use of the Hungarian algorithm (Allodi et al., 2016; Chen et al., 2021; Daniłowicz, 2020; Duan & Li, 2021; Li et al., 2017b; Lipovits et al., 2021; Meneses et al., 2020; Piao et al., 2016; Riahi & Bilodeau, 2015; Wang et al., 2021; Wu et al., 2019; Wu & Li, 2016; Yu et al., 2020; Zhang et al., 2020), LSTM based data association (Chang et al., 2020; Farhodov et al., 2020; Kim et al., 2021; Pang et al., 2021; Tan et al., 2018; Wan et al., 2018; Weng et al., 2020; Xu & Zhou, 2018), tracklet association (Bae & Yoon, 2017; He et al., 2020a; Wang et al., 2016; Yang et al., 2018a) and greedy data association methods (Kim et al., 2021; Singh et al., 2017, 2017, 2017). The incorporation of the Hungarian algorithm within current approaches is still the most common method. A two-step approach deduced into a compact data association method was proposed by Piao et al. (2016) to improve the simplicity and robustness by splitting the association tasks between a low and high-level stage. The results show more accuracy with less noisy scenes which may require a more rich appearance model. To handle the merging of obstacles, the Hungarian algorithm was applied by Allodi et al. (2016) which improved processing time and robustness but required improvement in maintaining tracks. A discriminative model applied to multi-person tracking was defined by Li et al. (2017b) to track through occlusion and accommodate long periods of time. On the other hand, Riahi and Bilodeau (2015) applied a generative appearance model to improve the robustness and accuracy but with a high number of identity switches. Applied with a zero-mean Gaussian function, Wu and Li (2016) defines a speed promoting method with a focus on occlusion handling. Wu et al. (2019) introduced a joint association matrix construction with the implementation of the Hungarian algorithm to predict observations and allow more successful tracks.

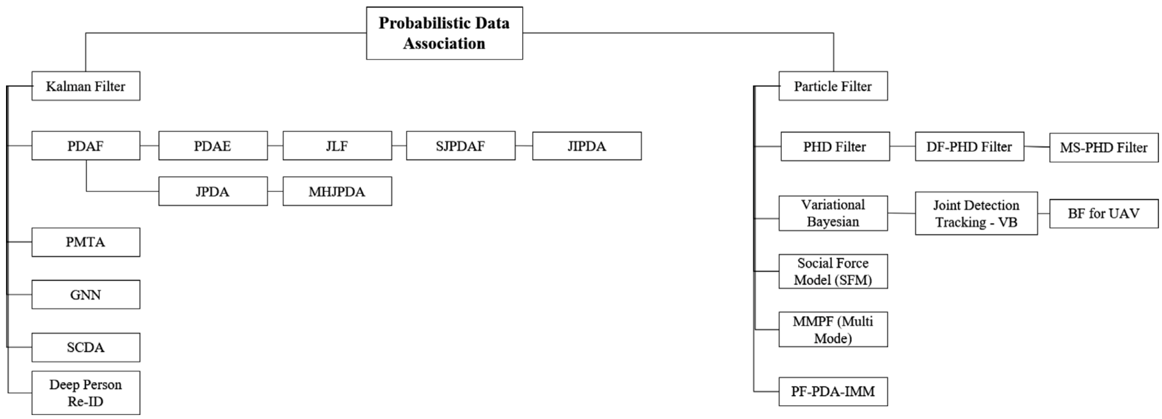


**Fig. 4.** A classification of Probabilistic Data Association methods applied within recent years. Firstly, we divide the techniques based on the parent model design by Kalman Filter or Particle Filter. The classifications are further detailed based on extensions of these models. On the left branch, Kalman Filter extends to PDAF versions of PDA as well as GNN and Deep Person Re-Identification. On the right side, the particle filter extends to PHD, Bayesian methods, Social Force and Multi-Mode models.

Singh et al. (2017) proposed a threshold-based greedy algorithm to find a locally optimal solution without the expense of computation complexity. The idea combines components of tracklet association to associate confirmed reliable tracklets in sets of pairs; however, while the objective was to ensure a high-speed performance of the algorithm, there was a compensation with qualitative results. A combination of part-based matching and linear programming with a greedy algorithm was described by Zhang et al. (2018) to offer a high track initiation rate. Considering the complex changes of human appearance through occlusion, illumination and pose changes, the method applies graph matching with a combination of affinity constraints. Performance results still display less satisfactory when handling occlusion in low-resolution videos. Another combination of the greedy algorithm using tracklet linking and particle filters was proposed by Jiang and Huynh (2017) to allow for reliable tracklet generation, occlusion handling and reduction of computation complexity. Considering its application in unconstrained visual sensor networks, the approach still requires improvement in the appearance and motion model.

With the word tracklets (Bewley, Ge et al., 2016) bound to define a set of certain confirmed tracks, methods of association by tracklets have been introduced (Sheng et al., 2018; Wang et al., 2018a, 2016). A tracklet confidence-based approach within a hierarchical association framework was proposed by Chen et al. (2018) for tracking multiple objects in airborne videos in order to add a level of robustness for complex video scenes. Experiments suggested a good appearance model with better results on appearance changes as well as to object re-identification, but this was hindered by long term occlusions and a higher false alarm rate. A similar approach which also applied the use of tracklet confidence was described by Bae and Yoon (2017) by pairing with a discriminative appearance learning model. The framework in Fig. 5 (Bae & Yoon, 2017) illustrates the levels of association based on high confidence (HC) and low confidence (LC). While offering accuracy and robustness, there is a limitation with this application to linear models only.

Apart from confidence measurement approaches (Bae & Yoon, 2017, 2017; Liu et al., 2018a), a tracklet affinity function association method by generalized linear assignment was proposed by Yang et al. (2018a) for the improvement of sparse representation and tracking consistency. Another tracklet association approach based on network flow optimization was described by Wu and Li (2016) aimed at preventing identity switches and recovering missing detections but as the number of targets increased in the frame, the computational load also increased exponentially.

While neural network deep learning methods have recently been popular in object detection and tracking, particularly with YOLO, SSD and Faster RCNN, it has also expanded to the data association framework with a focus on developing the temporal state. Particularly due to the strong memory component of LSTM (long short-term memory) modules as well as the effectiveness with non-linear transformation, LSTM-based data association has been applied to linear assignment situations as an extension described by Liu et al. (2019) focusing on the modules temporal functionality for assignment prediction by one target at a time. A sequence of multiple LSTM networks is designed for a step-by-step prediction of probability and measurement. Performance comparison results demonstrated a more robust and improved performance time when compared to legacy methods — the Hungarian Algorithm and JPDA. Zhang et al. (2019) also applied an LSTM based method in concatenation with a JPDA-RNN to tackle the issue of dense clutter; however, the paper highlights better performance needs to be researched in real-time situations due to a trend a delayed initiation and termination identified during online tracking. Another demonstration of the LSTM application to data association in deep neural networks by Yao et al. (2018) focused on the problem of initiation and termination of object trajectories and noisy output with particle detectors in the aspect of particle tracking. A frame-by-frame application of a temporal sliding window is applied using an LSTM layer in the form of an RNN with the results identifying the ability to capture intricate motion patterns that are more difficult to identify with traditional models, though currently the model is only based on a two-frame assignment approach and needs to be developed for multi-frame approaches.

Considering accuracy (Babaee et al., 2018) and stability (Liu et al., 2018b) as an optimal vantage point, a method of improving the temporal aspect of object tracking was defined through the use of LSTM's (Babaee et al., 2018; Farazi & Behnke, 2017; Xiang et al., 2019) to handle situations like long term occlusions. A combination of LSTM stacks with YOLO detected objects introduced by Tan et al. (2018) aided in improving the temporal relationship between frames. By using the addition of Euclidean distance to calculate the measurement between objects, the framework allowed for good adaptation to appearance changes, but the complexity of the network decided the computational efficiency. Paired with a region-based appearance model, an LSTM based pose model was illustrated by Xu and Zhou (2018) to measure the similarity between different identities. The LSTM layer allowed for improved speed with deep feature extraction, but the model improves tracking accuracy by adding more association cues. The model also splits between a hard and easy association step illustrated in Fig. 6 (Xu & Zhou, 2018)

A Siamese LSTM network proposed by Wan et al. (2018) established a model for improved state estimation and allowed for interpretation of both temporal and spatial components non-linearly.

New hierarchical merged techniques aimed at innovation include a visual appearance affinity model described by Bewley, Ott et al. (2016) to improve sequential adaptation. The advantage of the model came

**Table 2**
Probabilistic Methods with a summary of the advantages and disadvantages for each method.

| Technique | Advantages | Disadvantages |
|---|---|---|
| Permutation Matrix Track Association (PMTA) (Lee, Kanzawa et al., 2018) | Stability, Robustness | Customized to a specific scenario |
| Multi Model Smooth Variable Structure Filter (MMSVSF) (Luo et al., 2019) | Accuracy, robustness | Need more dynamic movements of surrounding objects |
| Global Nearest Neighbour (GNN) for Field-of-View Sensors (Cabrera, 2018) | Low number of lost tracks | Performance reduces over more objects entering the experiment |
| Social Force Model (SFM) within Markov Chain Mote Carlo (MCMC) (Feng et al., 2016) | Improved localization and cardinality | Computational complexity |
| Structural Constraints Data Association (Yoon et al., 2019) | Validated under unexpected camera motion, reduction in mis-detections and false positives | Less focus on temporal constraints like limited video lengths |
| Multi Mode Particle Filter (PF) Data Association (Ritter et al., 2018) | Improved accuracy in detection and tracking | No method to estimate bias, enhancement in tracking accuracy is small |
| PDAE - elliptical (Godinez & Rohr, 2014) | Robust to errors arising from spot detection, operates well on low samples | Performance is average on low SNR levels, computational time |
| Joint Detection and Tracking - Variational Bayesian (JDT-VB) (Lan et al., 2016) | Improved state estimation | No simulation verification |
| Joint Likelihood Filter (Rasmussen & Hager, 2001) | Improvement in handling occlusion | Confusion in classification of noise for some scenarios |
| Association Log Likelihood (Altendorfer & Wirkert, 2016) | Performs well in steady-state scenarios | No advantage in predicted track co variance matrices of arbitrary shape |
| Diffusion Particle PHD Filter (D-PPHDF)/ Multi Sensor Particle PHD Filter (MS-PPHDF) (Leonard & Zoubir, 2019) | Speed of accuracy | Poor in densely populated scene |
| Bayesian Filtering for Unmanned Aerial Vehicles (UAV) (Barkley & Paley, 2017) | Maintains tracker accuracy | Average accuracy for motion model of targets |
| Sample Based JPDAF (SJPDAF) (Zhang & Tang, 2018) | Accuracy improvement with inter-object occlusion | Computational complexity |
| Markov Chain Data Association with Joint Integrated PDA (JIPDA) (Lee et al., 2017b) | Reduced computational time, improved accuracy in dense clutter | Average results with false track discrimination and target retention |
| Poisson Spatial Measurement Model for JPDA (Yang et al., 2018b) | Requires no clustering and partitioning | Number of target tracks is fixed |
| Likelihood Based Data Association with Sampling Methods (Granström, Svensson et al., 2017) | Improved tracking performance | Complexity with implementation efficiency |
| JPDA for extended target tracking (ETT) (Vivone et al., 2015) | High tracking accuracy, limited false alarm rate, support real time requirements | Results consider the complete MOT framework with marine radar data |
| Improved JPDA (Hamid Rezatofighi et al., 2015) | Reduced computational complexity, good performance with noisy detections and occlusion | Average speed and not generalized |
| Multiple Hypothesis JPDA (MH-JPDA) with fixed interval (Stauch et al., 2017) | Tracking precision and accuracy with close spaced objects, real-time functionality | Initial ambiguity |
| Tracklet level association in MHT (Sheng et al., 2018) | Computational efficiency in solving MWIS problems with MHT | No specific improvement in tracking performance compared to existing algorithms |
| PDAF with compound segmentation technique (Mahemuti et al., 2016) | Reasonable tracking accuracy with low density microtubule video | Incorrect tracking estimation with complicated compound objects |
| JPDA for automated multi target tracking (Hunde & Ayalew, 2018) | High confidence measurement data association can improve tracking performance | Reduced performance in more complex traffic scenarios |
| PF-PDA-IMM scheme (Chalvatzaki et al., 2018b) | Improved accuracy, robustness to occlusion | Requires development for more complex motion models |
| JPDA in trajectory optimization (He et al., 2020b) | Single scan approach that reduces computational burden | Accuracy reduces as the count of modality increases |
| MHT in Graph neural networks (Rangesh et al., 2021) | Capable of working in real time | A difficult estimation problem with lower particles |
| JPDA with reinforcement learning (Qu et al., 2020) | Shorter execution time and effective in dense clutter. | Temporal range is set to a limited time space |

from the ability to learn without explicit labelling; however, it currently runs as offline based only and generates some track fragmentation. A confidence score based appearance model in a hierarchical data association framework to accommodate online multiple object tracking proposed by Liu et al. (2018a) allowed for faster processing time and accuracy with comparable but not improved results on the MOT datasets. A Part-based association technique was described by Jiang et al. (2017) to improve robustness based on appearance, position and motion, but there is the failure to capture temporal information effectively. Stochastic progressive association across multiple frames coined SPAAM was introduced by Elliethy and Sharma (2018) which had the advantage of accuracy with computational efficiency but improvement in the learning of richer spatial information was required. An increased usage of the Mahalanobis distance paired with the Hungarian algorithm has is apparent since 2020 in Salscheider (2021, 2021) and Wang et al. (2020b) confirming a common improvement in accuracy.

In a qualitative aspect, the hierarchical methods are more spread out compared to probabilistic methods with more techniques offering accuracy, robustness, speed ad offer a door into more deep learning architectures. A particular drawback with hierarchical methods is the computational processing complexity of the network added when multiplying the number of layers and splitting tasks between multiple association levels. Fig. 7 illustrates the taxonomy breakdown of the hierarchical methods mentioned (see Table 3).

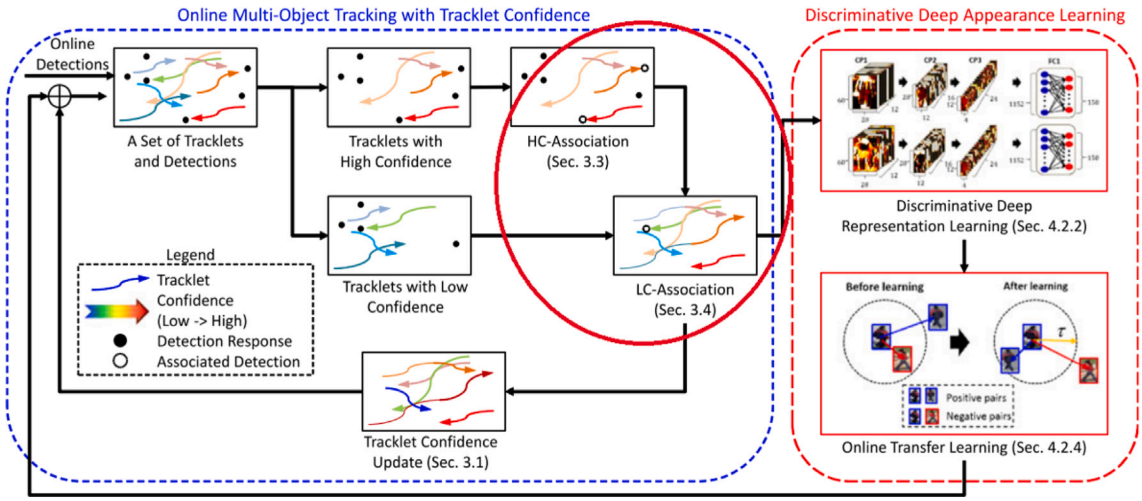


**Fig. 5.** A proposed framework for Confidence Based Data Association based on High Confidence and Low Confidence Association. These tasks are performed before appearance learning through deep discriminative methods. The threshold was configured with High Confidence being labelled for values over 0.5, however, the paper suggests that tracking performance was not affected by this threshold (Bae & Yoon, 2017).

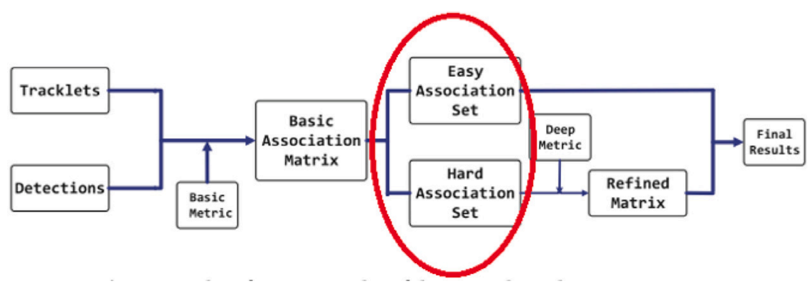


**Fig. 6.** A High-level flow for Easy Association and Hard Association. At first, a basic affinity matrix based on a basic association metric which then divides into easy and hard associations according to its basic score. The Hungarian algorithm is then applied to get a final association value (Xu & Zhou, 2018).

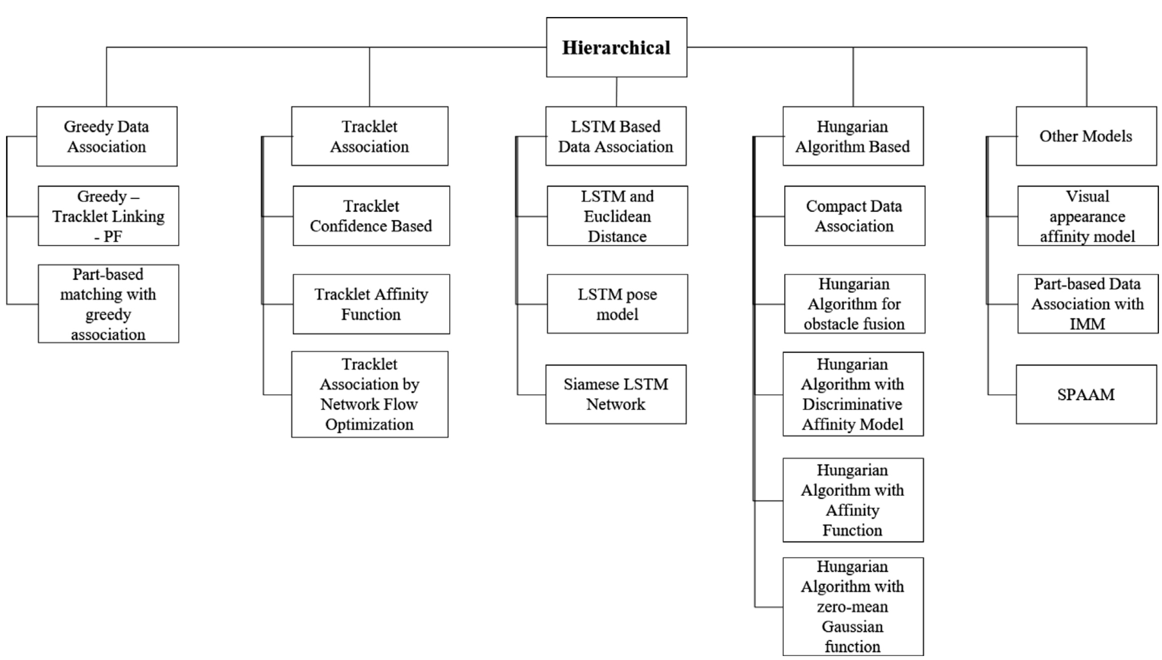


**Fig. 7.** Breakdown of Hierarchical Data Association Methods which start with more parent classifications such as Greedy Association, Tracklet Association, LSTM Based, Hungarian Algorithm based and other models. The application of multi-layer approaches allows for more research into hybrid approaches which merge two types of methods to obtain better results. Illustrated by the branch lengths, it is still popular among the use of the Hungarian Algorithm including combinations of the Hungarian Algorithm with other functions.

**Table 3**
Hierarchical Methods with a summary of the advantages and disadvantages for each method.

| Technique | Advantages | Disadvantages |
|---|---|---|
| Greedy data association with tracklet linking and particle filters (Singh et al., 2017) | Focus on high speed performance | Slight decrease in qualitative performance |
| Hierarchical data association tracking based on tracklet confidence (Chen et al., 2018) | Better single target tracking with appearance change and object re-identification | High false alarm rate of detections and poor re-identification after long term occlusion |
| LSTM and Euclidean distance (Tan et al., 2018) | Good adaptation to appearance changes of objects, fast performance | Average computational efficiency |
| Visual appearance affinity model (Bewley, Ott et al., 2016) | Model can learn without explicit labelling | Offline based only, some track fragmentation |
| Compact data association (two level association with assignment by Hungarian algorithm) (Piao et al., 2016) | Computationally efficient, robust | Accuracy is high based on less noisy level results, requires a better appearance model |
| Tracklet confidence-based data association (Bae & Yoon, 2017) | Accuracy, robustness | Only applies to linear motions |
| Part-based matching, linear programming and greedy data association (Zhang et al., 2018) | High track initiation rate | Poor results on occlusion and low resolution video |
| Greedy data association with tracklet linking and particle filters (Jiang & Huynh, 2017) | Reliable tracklet generation, occlusion handling, linear computational complexity | Average robustness, requires appearance and motion model evaluations |
| LSTM pose model with region-based appearance model (Xu & Zhou, 2018) | Speed with deep feature extraction, occlusion handling | Requires more tracking accuracy by adding more association cues |
| Hungarian algorithm for obstacle fusion and tracking association (Allodi et al., 2016) | Robust, fast processing time | Need improvement in maintaining tracks |
| Siamese LSTM network (Wan et al., 2018) | Accuracy | Computational complexity |
| Discriminative affinity model with Hungarian algorithm (Li et al., 2017b) | Tracks through occlusion and long periods of time | Requires more generalization, accuracy scores are comparable but not significantly improved |
| Affinity function with Hungarian algorithm (Riahi & Bilodeau, 2015) | Robust, high accuracy | High number of ID switches |
| Tracklet affinity function association by generalized linear assignment (Yang et al., 2018a) | Improved tracking consistency | Low computational efficiency |
| Confidence score-based appearance model (Liu et al., 2018a) | Accuracy, fast processing time, robustness | Results are comparable on MOT datasets |
| Part-based data association with IMM tracking (Jiang et al., 2017) | Robustness | Does not capture temporal information effectively |
| Hungarian algorithm with zero-mean Gaussian function (Wu & Li, 2016) | Speed promotion | Missing detections lead to false tracking alarms |
| Tracklet association by network flow optimization (Wang et al., 2016) | Prevents identity switches and recover missing detections | As targets increase, computational speed decreases, occlusion hinders performance |
| Stochastic Progressive Association Across Multiple Frames (SPAAM) (Elliethy & Sharma, 2018) | Accuracy, computational efficiency | Challenge with limited spatial details |
| LSTM-based aggregated mode (Chang et al., 2020) | Improved prediction over LSTM | some environmental factors make it difficult to verify testing values |
| Hungarian algorithm in R-CNN (Daniłowicz, 2020) | Used in real time situations | Limited information about dealing with temporal information |
| Hungarian algorithm for the assignment problem (Meneses et al., 2020) | improved tracking accuracy and efficient computation | – IOU and SORT trackers compare a faster identity switch rate |
| LSTM in a Graph Neural Network (Weng et al., 2020) | improves discriminative feature learning | limited information on speed |
| Mahalanobis distance and Hungarian algorithm (Salscheider, 2021) | reduced miss rate and localization error | requires some computational time |
| Adaptive fusion model based on kalman filtering and LSTM (Wang et al., 2021) | accuracy and speed improvement | limited specific information on robustness |
| Hungarian algorithm applied in an LSTM framework (Yu et al., 2020) | Shows stability in both long-term and short-term | Some noise such as fog creates lesser performance |
| Bilinear LSTM and greedy association (Kim et al., 2021) | Achieves real time tracking performance | Limited information on comparison of computational cost |
| LSTM with tracking association (Farhodov et al., 2020) | Supports long term tracking in real time and can work in online mode | Limited information on comparison of computational cost |

### 3.4. IMM, Kalman filter and Fuzzy data association

Areas of interest covered to solve the data association problem in multiple object tracking include the use of Interactive Multiple Models (IMM (Faber et al., 2018; Hu et al., 2020; Kulmon & Stukovska, 2018; Wu & Hong, 2005), the original Kalman Filter (KF) (Jeong et al., 2014; Sahbani & Adiprawita, 2016) and Fuzzy data association (Li et al., 2017a; Liu et al., 2020b; Raboaca et al., 2020; Tafti & Sadati, 2010; Xi-yang et al., 2018). A combination of the Kalman filter and the Mahalnobis distance has been applied frequently in recent MOT frameworks such as Cantas et al. (2021), Huang et al. (2020), Li et al. (2020, 2020) and Wang et al. (2020a). To allow a framework to be more adaptable for online tracking, Xi-yang et al. (2018) introduced MOANA, formulated through time and space, which applies a tracking mode through Kalman filtering. The model is applicable to the adaptation of 3D properties. While the model offers a more robust and efficient solution, it currently only works with static camera footage. Also applying properties through the extended version of the Kalman filter with a bipartite graph model method and Hungarian algorithm hybrid, Mei et al. (2017) was able to perform multiple target tracking in real-time but this only applied with the extraction of 2D target information. A multi-view approach for 3D reconstruction applied an MCMC based approach with tracklet association described by Tang et al. (2018a) to offer more efficiency in multi-view tracking, but less information is provided about computational efficiency. In a similar environment of multiple pedestrian tracking, an Interactive Multiple Model (IMM) was combined with the Munkres algorithm to offer a more robust solution by Jiang and Huynh (2017), but the performance diminishes as the scene becomes crowded. Another IMM based tracking association method proposed by Yuan et al. (2017a) combined with Multiple Hypothesis Testing (MHT) was called IMM-SMHT (Interacting Multiple Model with a sequential multiple hypothesis test model and

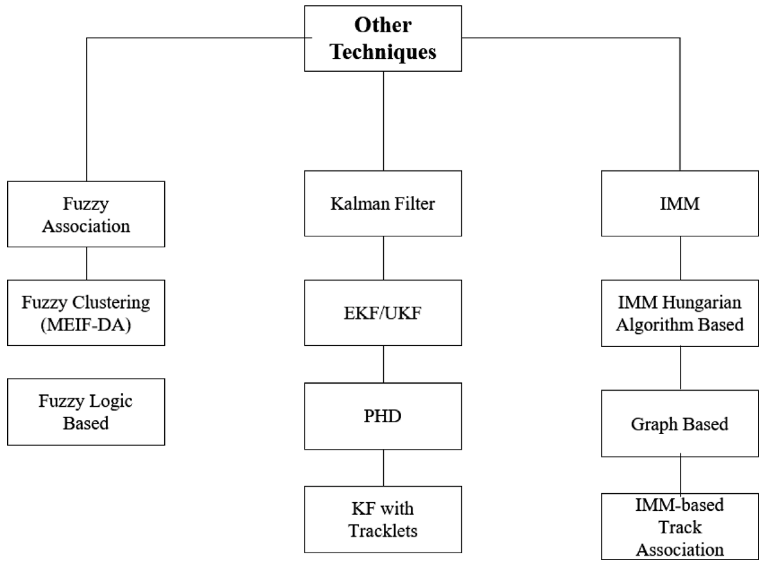


**Fig. 8.** Breakdown of other techniques — IMM, Kalman Filter, Fuzzy Association. These include methods which have had some but little presence in recent literature, so they have been grouped for the purpose of this survey. Though Kalman Filter is a category under the Probabilistic approach, here we identify models that take advantage of Kalman filtering without using the PDAF approaches.

also offered robustness with added high accuracy but still limited by a computational cost. Considering the multi-view issue, Wang et al. (2014) used appearance modelling to for the track association task using a competitive major feature histogram (CMFH) which is employed for addressing the problem of disjoint camera views. Wu and Hong (2005) suggested a combination of the IMM method with Probabilistic data association to process long video segments and show a better performance over the Kalman Filter.

A different technique involves using Fuzzy data association to calculate fuzzy association probabilities by design of a Fuzzy Inference System (FIS) based on affinities of appearance and motion. An approach by Li et al. (2017a) offered good performance results while Xi-yang et al. (2018) fused a Kalman filter with maximum entropy intuitionistic fuzzy data association which becomes a reconstructed intuitionistic model. Though this method offers good robustness and accuracy results, the limit of computational cost is not removed.

An older method that still requires a mention is the graph-based approach (De Sousa & Kropatsch, 2015; Salvi et al., 2013). More popular in the late 2000s, it has notably been less often used in the last five years. Gulati et al. (2017) performed a comparison between Random Finite Sets (RFS) and Factor graphs for multi-sensor corporative localization to allow for better state estimations. According to the paper, factor graphs actually avoid the task of data association altogether. Factor graphs display a higher execution time due to offline batch optimization, but both methods offer improved scalability. Recent approaches of Graph based methods include a graph similarity model that focuses on improving robustness and accuracy (Liu et al., 2020a) and the use of neighbour graphs with the Hungarian algorithm (Liang et al., 2020). An illustration for the taxonomy of other techniques such as IMM, Kalman Filter-based and Fuzzy data association is shown in Fig. 8 (see Table 4).

### 3.5. New techniques

By grouping some new techniques together upon qualitative analysis, there is more development on accuracy results with a focus on improving the computational complexity (Elliethy & Sharma, 2018; Kokul et al., 2015), however, the basis of accuracy is applied to network training for specific scenarios and discarding the generalization ability. Hybrid data association techniques to apply the best of both local and global association features have been proposed by Yang et al. (2017) who determines association as a minimum cost multi-commodity flow problem to handle online data and scenarios where the number of objects is large. With a goal to improve performance and overall accuracy for tracking, the framework did require more definition with the motion and shape cues to allow for greater stability. In terms of a minimum cost multi-way solution, Park et al. (2014) applied this data association method with a Lagrange Dual solution to optimize tracking between interacting objects. This allowed a model with better generalization ability along with accuracy but still adding some computational complexity. Another hybrid proposal by Dai et al. (2018) applied an affinity model with the detection-detection association where tracklets are created based on the detections from individual frames. This created a better accuracy reading in complex scenarios, but there is still future work to improve the object interaction model and shape information models.

Particular challenges in recent years that have gained research into improving the performance for multiple object tracking involve crowded scenes of pedestrians or traffic. Dehghan and Shah (2017) formulated this problem as binary quadratic programming and applied Frank Wolfe optimization for the association task. The method was combined with SWAP steps to reduce the computational complexity, a necessary improvement when the number of objects in the scene or frame increases over time. With performance and accuracy at the forefront of improvement, there is a toll taken on computational complexity. An analysis performed by Niedfeldt et al. (2017) applied a recursive RANSAC approach over a sequential method to allow for improved parameter estimation over multiple signals and to offer better confidence of convergence to the correct solution. The recursive method provides more simplicity and tracking continuity performance; however, the parameters are environmentally dependent, which makes the model poor for generalization.

A simple improvement on robustness came from Steyer et al. (2018) who experimented on a new particle labelling association method which made use of low-level particle representation in space defined by grid cells and object tracks. Speed and computational cost are discussed less extensively in the paper. Fig. 9 (Steyer et al., 2018) is an illustration of the particle labelling association within defined grid cells and the phases of particle labelling for correct tracking.

In the application of a single camera and inter-camera dataset, Tang et al. (2018b) applied a bottom-up clustering strategy paired with a loss function resulting in enhanced robustness and speed improvement.

**Table 4**
Other Methods with a summary of the advantages and disadvantages for each method.

| Technique | Advantages | Disadvantages |
|---|---|---|
| Kalman MOANA : An Online Learned Adaptive Appearance Model for Robust Multiple Object Tracking in 3D (Xi-yang et al., 2018) | More robust and efficient | Only works on static cameras, future work on moving cameras |
| Bipartite graph model based method with Hungarian algorithm (Mei et al., 2017) | Tracks multiple targets in real time | Extracts only 2D target information, future work on 3D features |
| Kalman filter for multi view object tracking by data association (MCMC based approach) (Tang et al., 2018a) | Efficiency in multi view tracking | Less information on computational efficiency |
| IMM IMM with The Munkres Algorithm (Jiang & Huynh, 2017) | Robust | Cannot handle more complex or crowded scenes |
| IMM-SMHT (IMM with sequential multiple hypothesis test) (Yuan et al., 2017a) | Robustness and high accuracy | Computational cost |
| Fuzzy Fuzzy logic data association (Li et al., 2017a) | Better performance | Little information on stability |
| Kalman filter with maximum entropy intuitionistic data association (Xi-yang et al., 2018) | Robustness and accuracy | No information on computational time |
| Factor Graphs vs RFS framework (Gulati et al., 2017) | Scalability improvement and bandwidth reduction for both methods | RFS approach degrades over an increasing number of sensors while factor graphs have a higher execution time |
| Kullback–leibler differential entropy equation-based measurement data association (Hu et al., 2020) | The model is more adaptable with increased noise | Computational cost increase |
| Kalman filter with the Mahalanobis distance to evaluate motion distance (Huang et al., 2020) | Reduces the time costs of data association using 3D integral image constraints | Little information on computational complexity |
| Self supervised approach to associating objects using Kalman filter and Mahalanobis distance (Wang et al., 2020a) | Useful suppression of incorrect associations during self supervised training | Unclear if human efficiency correlates to the computational efficiency of the model |
| Data association between perception and V2V communication sensors (Cantas et al., 2021) | Method can be effective for curved roads and intersections | More tests need to be performed in complex scenarios and crowded scenes |
| Extended Kalman filter using IP (innovation projections) (Joerger & Hassani, 2020) | Lower level of wrong associations | Complex input parameters with Lidar |
| A priority data association policy for multitarget tracking (Zeng et al., 2020) | Improve performance for risk assessment | Performance diminishes with multi target tracking |
| Two stage fuzzy logic association integrated with Kalman filtering (Liu et al., 2020b) | Improve performance with multiple targets and complex movements | Sometimes false sizes of targets are generated due to reflections or disturbance |
| GSM:graph similarity model for multi-object tracking (Liu et al., 2020a) | Focuses on improving robustness and accuracy in association | Comparisons are only made with online settings for simplicity |
| Neighbour graph with GCN (graph convolutional networks) (Liang et al., 2020) | Applies features of the spatio-temporal domains based on neighbour selection | Requires more research to improve object detection |

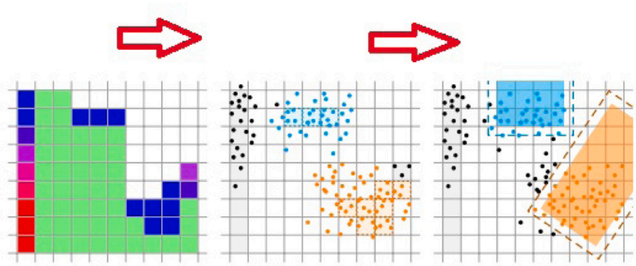

**Fig. 9.** The figure illustrates the particle labelling association concept using grid cells. This is further described in the paper by Steyer et al. based on the identification of particle labels to define bounding boxes and updating tracks (Steyer et al., 2018).

Based on a similarity function to be applied on videos from non-overlapping cameras, Choi and Jeon (2016) focused on accuracy with re-identification, but the description limits the results on computational cost and robustness. In a different problem set with partial views, Wong et al. (2015) employed a clustering-based data association method as the Dirichlet Process Mixture Model (DPMM) given multiple camera viewpoints. Given improved results in computational time and estimation, there was an identification of degradation over longer temporal states. Combined with sparse representations, a multi-frame data association using an energy minimization procedure described by Fagot-Bouquet et al. (2016) offered a robust approach with fewer association errors but at a computational cost. Basing an association decision off of a confidence score is not a new technique but the application in a Partial Least Square method (PLS) described by Lee, Kim et al. (2018) improved the evaluation for the confidence of a tracklet in terms of length and continuity. Though there is an overall improvement in the performance of the tracker, there is no specific discrimination detailing the data association aspect as a factor for the improvement. Rather than a cost function, Jaiswal et al. employed a gain function which was defined based on the position of the object and the image intensity to allow for more accuracy but limited to specific tasks.

To work in a scene with unknown backgrounds, Punchihewa et al. (2018) proposed a GLMB Joint Object Clutter Model, a filter extracted from Random Finite Sets (RFS) whose parameters evolve with time. The results lead to an improved performance rate, but the complexity increases as the scene become crowded.

Adding the problem of data association into the realm of deep learning has still been met with some reluctance. A Siamese CNN approach initiated by Lean-Taixe et al. adds the association task within the CNN layers by including a patch similarity in the form of a cost function. While the use of CNN enhances the results in speed, there is more development required to improve accuracy in rich situations. Adding to accuracy improvement included an on detection data association and segmentation approach by Tian et al. (2018) which overall did not match up to state-of-the-art performances, an enhanced identity association by Gan et al. (2018) which generated some confusion during occlusion and interaction with other targets, and a stochastic optimization method by Granström, Renter et al. (2017) which did not offer enough information on robustness and generalization (see Fig. 11, Tables 5–9).

**Table 5**
New Methods with a summary of the advantages and disadvantages for each method.

| Technique | Advantages | Disadvantages |
|---|---|---|
| Hybrid data association as a min-cost multi commodity flow problem (Yang et al., 2017) | Improved performance for online tracking, accuracy | Need more definition on motion and shape cues |
| Modified Frank Wolfe algorithm with SWAP steps (Dehghan & Shah, 2017) | Reduce computational complexity | Run time increases as people increases |
| Recursive RANSAC (Niedfeldt et al., 2017) | Simple to implement, efficient, robust | Computational complexity |
| Particle Labelling Association for grid-based object tracking (Steyer et al., 2018) | Robust | Less information on speed and computation cost |
| Similarity function based association for non overlapping cameras (Choi & Jeon, 2016) | Accuracy on re-identification | Less information on robustness, speed and computation |
| Dirichlet Process Mixture Models (DPMM) - A clustering based data association (Wong et al., 2015) | Improved estimation and computational time | Results degrade over longer temporal states |
| Multi-frame data association with sparse representation (Fagot-Bouquet et al., 2016) | Robust, less association errors | Computational cost |
| Hybrid data association with affinity and detection-detection association (Dai et al., 2018) | Accuracy in complex sequences | Need richer interaction model and shape information model |
| CNN based data association (Leal-Taixé et al., 2016) | Efficiency in speed | Need more accuracy in complex situations |
| Confidence based data association with Partial Least Squares (PLS) method. (Lee, Kim et al., 2018) | Improved overall performance | The tracking results may not depend on the association aspect in the paper |
| Minimum cost multi-way data association with Lagrange Dual solution (Park et al., 2014) | Accuracy and generalization ability | Convergence with polynomial time, some computational complexity |
| Gain functions for data association (Jaiswal et al., 2018) | Accuracy | No generalization yet - applies only to a specific trained task |
| GLMB Joint Object Clutter model (Punchihewa et al., 2018) | Improved performance rate | More complexity as the scene becomes crowded |
| On detection data association and segmentation (Tian et al., 2018) | Accurate segmentation results | Poor performance with overall comparison |
| Enhanced identity association (Gan et al., 2018) | High accuracy and precision with low ID switches | Confusion by occlusion and interaction with other targets, not generalized. |
| Stochastic Optimization (Granström, Renter et al., 2017) | Accuracy, computationally efficient | Less information on robustness and generalization |
| Bottom up clustering strategy with loss function (Tang et al., 2018b) | Enhanced robustness, speed improvement and re-identification | Computational complexity |
| Tracklets (long term and short term) (Yang et al., 2020) | Learns better appearance features for more effective association | Cannot be used on real time situations |
| Attention with transformer network data association (Hung et al., 2020) | Improves results with occlusion | Requires improvement with spatio-temporal dependancies |
| TransTrack, Box association with Kuhn–Munkres (KM) algorithm (Sun et al., 2020) | Improves accuracy | Limited information about long term tracking |
| Attention association with TrackFormer (Meinhardt et al., 2021) | Attention operations improve appearance modelling and improves segmentation | Limited information about computational cost |
| DASOT: a unified framework integrating data association (Chu et al., 2020) | Focuses on improving computational cost through learning discriminative features | Limited information on adaptability to long term tracking |
| Data association using geometry priors (Chen et al., 2020) | Generates a combined cost matric to improve accuracy | Requires further study for the inclusion of more sensor information |
| Embedding association using attention (Guo et al., 2021) | Improves robustness and tracking performance | Requires further study for computational efficiency |

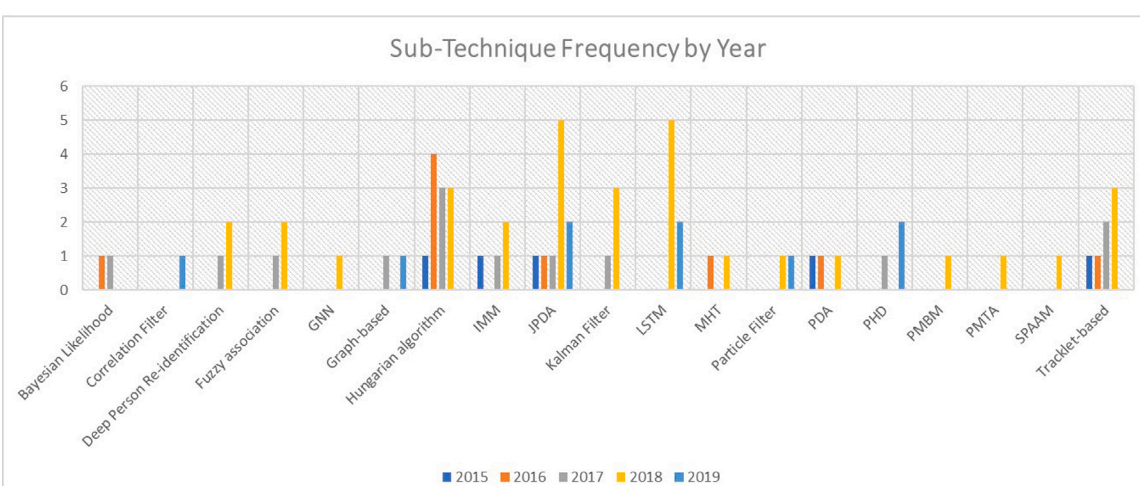


**Fig. 10.** The bar graph illustrates the sub techniques used in recent publications and colour coded by year ranging from 2015 to the most recent papers published in 2021. It further provides a better distinction of the more prominent methods utilized in the current research area of multiple object tracking — particularly the Hungarian algorithm, JPDA and LSTM modules displaying greater popularity. An association method which also demonstrates research growth in applications over the last four years is tracklet-based data association solutions and the integration of the Kalman filter.

**Table 6**
A Qualitative analytical summary of Probabilistic methods based on stability, accuracy of results described in the publication, robustness, cardinality and speed based on computational complexity.

| Technique | Stability | Accuracy | Robust | Cardinality | Speed |
|---|---|---|---|---|---|
| Permutation Matrix Track Association (PMTA) (Lee, Kanzawa et al., 2018) | Yes | | Yes | | |
| Multi Model Smooth Variable Structure Filter (MMSVSF) (Luo et al., 2019) | | Yes | Yes | | |
| Global Nearest Neighbour (GNN) for Field-of-View Sensors (Cabrera, 2018) | | Yes | | | |
| Social Force Model (SFM) within Markov Chain Mote Carlo (MCMC) (Feng et al., 2016) | | | | Yes | |
| Structural Constraints Data Association (Yoon et al., 2019) | | Yes | | | |
| Multi Mode Particle Filter (PF) Data Association (Ritter et al., 2018) | | Yes | | | |
| PDAE - elliptical (Ritter et al., 2018) | | | Yes | | |
| Joint Detection and Tracking - Variational Bayesian (JDT-VB) (Lan et al., 2016) | | Yes | | | |
| Joint Likelihood Filter (Rasmussen & Hager, 2001) | | Yes | | | |
| Association Log Likelihood (Altendorfer & Wirkert, 2016) | Yes | | | | |
| Diffusion Particle PHD Filter (D-PPHDF)/ Multi Sensor Particle PHD Filter (MS-PPHDF) (Leonard & Zoubir, 2019) | | Yes | | | Yes |
| Bayesian Filtering for Unmanned Aerial Vehicles (UAV) (Barkley & Paley, 2017) | | Yes | | | |
| Sample Based JPDAF (SJPDAF) (Zhang & Tang, 2018) | | Yes | | | |
| Markov Chain Data Association with Joint Integrated PDA (JIPDA) (Lee et al., 2017b) | | Yes | | | Yes |
| Poisson Spatial Measurement Model for JPDA (Yang et al., 2018b) | | | | | Yes |
| Likelihood Based Data Association with Sampling Methods (Granström, Svensson et al., 2017) | | Yes | | | |
| JPDA for extended target tracking (ETT) (Vivone et al., 2015) | | Yes | | | Yes |
| Improved JPDA (Hamid Rezatofighi et al., 2015) | Yes | | | | Yes |
| Multiple Hypothesis JPDA (MH-JPDA) with fixed interval (Stauch et al., 2017) | | Yes | | | Yes |
| Tracklet level association in MHT (Sheng et al., 2018) | | | | | Yes |
| PDAF with compound segmentation technique (Mahemuti et al., 2016) | | Yes | | | |
| JPDA for automated multi target tracking (Hunde & Ayalew, 2018) | | Yes | | | |
| PF-PDA-IMM scheme (Chalvatzaki et al., 2018b) | | Yes | Yes | | |
| JPDA with tensor decomposition (Krishnaswamy & Kumar, 2018) | | | | | Yes |
| JPDA with trajectory optimization (He et al., 2020b) | Yes | | | Yes | |
| MHT in Graph neural network (Rangesh et al., 2021) | Yes | | | Yes | |
| JPDA with reinforcement learning (Qu et al., 2020) | Yes | Yes | | | |

**Table 7**
A Qualitative analytical summary of Hierarchical methods based on stability, accuracy of results described in the publication, robustness, speed of results and computational complexity.

| Technique | Stability | Accuracy | Robust | Speed | Comp. Efficiency |
|---|---|---|---|---|---|
| Greedy data association with tracklet linking and particle filters (Singh et al., 2017) | | | | Yes | |
| Hierarchical data association tracking based on tracklet confidence (Chen et al., 2018) | | Yes | | | |
| LSTM and Euclidean distance (Tan et al., 2018) | | Yes | | Yes | |
| Visual appearance affinity model (Bewley, Ott et al., 2016) | | | Yes | | |
| Compact data association (two level association with assignment by Hungarian algorithm) (Piao et al., 2016) | | | Yes | | Yes |
| Tracklet confidence-based data association (Bae & Yoon, 2017) | | Yes | Yes | | |
| Part-based matching, linear programming and greedy data association (Zhang et al., 2018) | | | | Yes | |
| Greedy data association with tracklet linking and particle filters (Jiang & Huynh, 2017) | | Yes | | Yes | |
| LSTM pose model with region-based appearance model (Xu & Zhou, 2018) | | | Yes | Yes | |
| Hungarian algorithm for obstacle fusion and tracking association (Allodi et al., 2016) | | | Yes | Yes | |
| Siamese LSTM network (Wan et al., 2018) | | Yes | | | |
| Discriminative affinity model with Hungarian algorithm (Li et al., 2017b) | Yes | | | | |
| LSTM-based aggregated mode (Chang et al., 2020) | | Yes | | | |
| Hungarian algorithm in R-CNN (Daniłowicz, 2020) | | Yes | | Yes | Yes |
| Hungarian algorithm for the assignment problem (Meneses et al., 2020) | | Yes | | Yes | Yes |
| LSTM in a Graph Neural Network (Weng et al., 2020) | | Yes | | | |
| Mahalanobis distance and Hungarian algorithm (Salscheider, 2021) | | Yes | | | |
| Adaptive fusion model based on kalman filtering and LSTM (Wang et al., 2021) | | Yes | | | Yes |
| Hungarian algorithm applied in an LSTM framework (Yu et al., 2020) | | Yes | Yes | | |
| Bilinear LSTM and greedy association (Kim et al., 2021) | Yes | Yes | | | |
| LSTM with tracking association (Farhodov et al., 2020) | Yes | Yes | | Yes | |

## 4. Discussion

Still being one of the more popular methods to apply in a data association task, Probabilistic methods are still being used and upgraded/ extended to suite the scenario or video environment. While older algorithms like Global Nearest Neighbour (GNN), Particle and Kalman Filters, and Multiple Hypothesis Testing (MHT) are still used in similarity comparison or distance calculation, more emphasis has been placed on the extension or improvement of the Joint Probabilistic Data Association Method resulting in greater performance accuracy and robustness. Though accuracy is a common factor of improvement, the probabilistic approaches have shown more options in improving computational complexity over other categories. This indicates an important element when considering a real-time performing framework in the Multiple Object Tracking scenario. As illustrated by Fig. 10, JPDA still proves to be the most widely applied probabilistic technique and currently matches upcoming deep learning applications of LSTM within new research.

Hierarchical association methods have also proven to be still a popular technique with many applications using the multi-level technique

**Table 8**
A Qualitative analytical summary of other methods based on stability, accuracy of results described in the publication, robustness, speed of results and computational complexity.

| Technique | Stability | Accuracy | Robust | Speed | Comp. Efficiency |
|---|---|---|---|---|---|
| MOANA : An Online Learned Adaptive Appearance Model for Robust Multiple Object Tracking in 3D (Xi-yang et al., 2018) | | | Yes | Yes | |
| Bipartite graph model based method with Hungarian algorithm (Mei et al., 2017) | | | | Yes | |
| Kalman filter for multi view object tracking by data association (MCMC based approach) (Tang et al., 2018a) | | | | Yes | |
| IMM with The Munkres Algorithm (Jiang & Huynh, 2017) | | | Yes | | |
| IMM-SMHT (IMM with sequential multiple hypothesis test) (Yuan et al., 2017a) | | Yes | Yes | | |
| Fuzzy logic data association (Li et al., 2017a) | | | | Yes | |
| Kalman filter with maximum entropy intuitionistic data association (Xi-yang et al., 2018) | | Yes | Yes | | |
| Factor Graphs vs RFS framework (Gulati et al., 2017) | Yes | | | | |
| Kullback–leibler differential entropy equation-based measurement data association (Hu et al., 2020) | Yes | Yes | Yes | | |
| Kalman filter with the Mahalanobis distance to evaluate motion distance (Huang et al., 2020) | | Yes | Yes | Yes | |
| Self supervised approach to associating objects using Kalman filter and Mahalanobis distance (Wang et al., 2020a) | | | Yes | Yes | |
| Data association between perception and V2V communication sensors (Cantas et al., 2021) | Yes | | Yes | | |
| Extended Kalman filter using IP (innovation projections) (Joerger & Hassani, 2020) | | Yes | | | |
| A priority data association policy for multitarget tracking (Zeng et al., 2020) | Yes | | | | |
| Two stage fuzzy logic association integrated with Kalman filtering (Liu et al., 2020b) | | Yes | | | |
| GSM:graph similarity model for multi-object tracking (Liu et al., 2020a) | | Yes | Yes | | |
| Neighbour graph with GCN (graph convolutional networks) (Liang et al., 2020) | | Yes | | | |

**Table 9**
A Qualitative analytical summary of new methods based on stability, accuracy of results described in the publication, robustness, speed of results and computational complexity.

| Technique | Stability | Accuracy | Robust | Speed | Comp. Efficiency |
|---|---|---|---|---|---|
| Hybrid data association as a min-cost multi commodity flow problem (Yang et al., 2017) | | Yes | | | |
| Modified Frank Wolfe algorithm with SWAP steps (Dehghan & Shah, 2017) | | | | | Yes |
| Recursive RANSAC (Niedfeldt et al., 2017) | | | Yes | Yes | |
| Particle Labelling Association for grid-based object tracking (Steyer et al., 2018) | | | Yes | | |
| Similarity function based association for non overlapping cameras (Choi & Jeon, 2016) | | Yes | | | |
| Dirichlet Process Mixture Models (DPMM) - A clustering based data association (Wong et al., 2015) | | Yes | | | Yes |
| Multi-frame data association with sparse representation (Fagot-Bouquet et al., 2016) | | Yes | Yes | | |
| Hybrid data association with affinity and detection-detection association (Dai et al., 2018) | | Yes | | | |
| CNN based data association (Leal-Taixé et al., 2016) | | | | | Yes |
| Confidence based data association with Partial Least Squares (PLS) method. (Lee, Kim et al., 2018) | | Yes | | Yes | |
| Minimum cost multi-way data association with Langrange Dual solution (Park et al., 2014) | | Yes | | | |
| Gain functions for data association (Jaiswal et al., 2018) | | Yes | | | |
| GLMB Joint Object Clutter model (Punchihewa et al., 2018) | | Yes | | Yes | |
| On detection data association and segmentation (Tian et al., 2018) | | Yes | | | |
| Enhanced identity association (Gan et al., 2018) | | Yes | | | |
| Stochastic Optimization (Granström, Renter et al., 2017) | | Yes | | | Yes |
| Bottom up clustering strategy with loss function (Tang et al., 2018b) | | | Yes | Yes | |
| Tracklets (long term and short term) (Yang et al., 2020) | | Yes | | | |
| Attention with transformer network data association (Hung et al., 2020) | Yes | Yes | Yes | | |
| TransTrack, Box association with Kuhn–Munkres (KM) algorithm (Sun et al., 2020) | | Yes | | | |
| Attention association with TrackFormer (Meinhardt et al., 2021) | | Yes | | | |
| DASOT: a unified framework integrating data association (Chu et al., 2020) | | Yes | Yes | Yes | Yes |
| Data association using geometry priors (Chen et al., 2020) | | Yes | | | |
| Embedding association using attention (Guo et al., 2021) | | Yes | Yes | | |

to fuse advantages from multiple historical techniques. Clearly, the use of the Hungarian algorithm is still prevalent in the similarity function task, and distance calculation also highlighted in Fig. 10. Compared to probabilistic methods, the positive improvements in performance is more evenly distributed between accuracy, robustness and speed but little improvement or less mention on computational complexity. A confidence score based appearance model (Yang et al., 2018a) checked the most performance boxes (accuracy, robustness and computational complexity) among all the method comparisons combined. A particularly notable element is the improvement in handling temporal information with hierarchical methods such as the use of LSTM's for improved performances in re-identification. The bar graph in Fig. 10 further demonstrates the growing popularity with the use of LSTM modules (particularly with the incorporation of the Hungarian algorithm) because of their robust ability to fit into a deep learning multiple object tracking framework and the strong ability to handle the temporal state.

For the case of other techniques spanning IMM, Kalman Filter and Fuzzy data association, results were more swayed to the improvement of robustness and speed performance with less to no specific information on the stability and computational efficiency. The results comparison does illustrate the use of Kalman Filter merged with other methods such as the Hungarian algorithm, or an adaptive appearance model can enhance the speed performance, which makes it a good

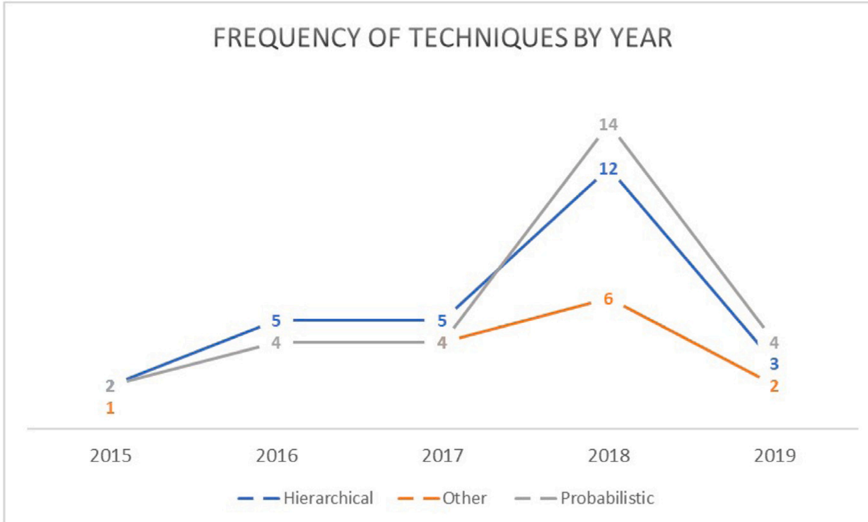


**Fig. 11.** The graph illustrates the count of papers categorized by year of publication and discussed under three major techniques discussed — Probabilistic, Hierarchical and Others. It provides a cumulative count for comparison purposes, simply indicating the high frequency between Hierarchical and Probabilistic methods of data association between 2015 and 2019. Patterns in recent research illustrate an increasing count of other techniques outside the traditional methods.

consideration when focusing on speed improvement for an MOT framework. Since 2020, there has been more use of the Kalman filter with a choice between the Mahalanobis distance or the Hungarian algorithm as a means of data association to improve computational efficiency.

The evaluation of new techniques is populated by hybrid approaches which try to incorporate best elements from left and right categories (for example, local and global associations). Maintaining the trend of previous techniques, there is still a heavy focus on the accuracy improvement results with almost all the new techniques reviewed showing some form of improvement in accuracy measurements compared to state-of-the-art performances. On another hand, there is no documented improvement on the stability of the method with some advancements showing in robustness, speed and computational efficiency.

As more research is being conducted to improve the computational speed of Multiple Object Tracking solutions, there is still a hole in the existing literature that covers a model solution which can verify performance improvements in all the measurement areas highlighted for this paper. Though accuracy is highlighted as the most significantly improved area within recent years, the uneven distribution of results shows that the balance of accuracy and speed still needs more work. For my future work, there are considerations to work on a hybrid framework that can best work with spatial and temporal information. To that end, using elements within the Hierarchical category such as enhancements to LSTM and gated solutions would be of key interest in regards to addressing both spatial and temporal parameters to support longer time associations.

## 5. Conclusion

In this paper, we compacted a summary and review of data association tasks proposed within recently published visual multiple object tracking frameworks and classified them under Probabilistic, Hierarchical, IMM, Kalman Filter, Fuzzy Association and New Technique based approaches. The advantages and disadvantages for each model were tabulated along with a performance comparison summary. The aim of this review was to organize models based on key technologies or procedures in order to perform better comparison evaluations. The detachment of the data association task for object tracking to be analysed and improved as an internal model confirms its importance in performance improvement. A breakdown of performance metrics such as the qualitative table analysis can give rise for future researchers to measure and compare the association task alone in a similar aspect. The review focused on algorithms applied particularly to the tracking of multiple objects in traffic surveillance video (either single or multi-camera based) and identified key weaknesses in appearance modelling, computational complexity and the struggle to enable a genuinely real-time performance.

Single Object Tracking has made leaps in improvement towards accuracy and real-time performance, but this has not completely translated well to the tracking of multiple objects. Research has shown a trend of using successful single object tracking methods and converting them into an MOT framework with the help of some data association algorithm to accurately maintain multiple tracks and detections simultaneously. The significance of this survey was to identify new techniques that have been developed and old techniques that have been modified to suit the task of successfully tracking multiple objects.

While some literature has covered reviews on the Multiple Object Tracking frameworks, our paper focuses on a subsection of the framework to better understand how we can improve separate components of a full framework to obtain better results. Since there already exist many reviews on the Object detection phase, there were still fewer reviews on data association in MOT itself previously, though few surveys were conducted with a limited reference list. Our paper tries to cover a wider analysis of recently published papers to identify a better indication of the MOT development direction.

Our qualitative analysis found a large number of publications still using Probabilistic and Hierarchical approaches due to their successful improvements in accuracy measurements and a more even distribution in positive results between stability, robustness and speed. A grouping of new techniques also demonstrated a focus on accuracy improvement results. In the near future, a trend is leaning towards a higher application of hierarchical methods due to its adaptability of hybrid approaches and its support for more long-term memory and long term association solutions.

While a key performance aspect for accuracy has been improved recently, it comes at the cost of having a situation or object-specific model, which becomes difficult to generalize. If the reader finds a particular model or technique interesting, they can refer to the referenced organization tables for information on the source papers.

## CRediT authorship contribution statement

**Lionel Rakai:** Conceptualization of this study, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Visualization, Writing – original draft. **Huansheng**

**Song:** Funding acquisition, Resources, Project administration, Supervision, Validation, Writing – review & editing. **ShiJie Sun:** Investigation, Project administration, Resources, Software, Visualization, Writing – review & editing. **Wentao Zhang:** Data curation, Visualization, Writing – review & editing. **Yanni Yang:** Data curation, Visualization, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This research was supported by the National Natural Science Foundation of China (Grant No. 62072053), the National Natural Science Foundation of China (Grant No. 62006026), and the China National Key R&D Program (Grant No. 2019YFB1600502).

## References

Al-Shakarji, N. M., Bunyak, F., Seetharaman, G., & Palaniappan, K. (2018). Multi-object tracking cascade with multi-step data association and occlusion handling. In *2018 15th IEEE international conference on advanced video and signal based surveillance* (AVSS), (pp. 1–6). IEEE.

Allodi, M., Broggi, A., Giaquinto, D., Patander, M., & Prioletti, A. (2016). Machine learning in tracking associations with stereo vision and lidar observations for an autonomous vehicle. In *2016 IEEE intelligent vehicles symposium (IV)* (pp. 648–653). IEEE.

Altendorfer, R., & Wirkert, S. (2016). Why the association log-likelihood distance should be used for measurement-to-track association. In *2016 IEEE intelligent vehicles symposium (IV)* (pp. 258–265). IEEE.

Anuj, L., & Krishna, M. G. (2017). Multiple camera based multiple object tracking under occlusion: A survey. In *2017 international conference on innovative mechanisms for industry applications* (ICIMIA), (pp. 432–437). IEEE.

Babaee, M., Athar, A., & Rigoll, G. (2018). Multiple people tracking using hierarchical deep tracklet re-identification. arXiv preprint arXiv:1811.04091.

Bae, S.-H., & Yoon, K.-J. (2017). Confidence-based data association and discriminative deep appearance learning for robust online multi-object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(3), 595–610.

Bar-Shalom, Y., Daum, F., & Huang, J. (2009). Estimation in the presence of measurement origin uncertainty. *IEEE Control Systems Magazine*, 82–100.

Barkley, B. E., & Paley, D. A. (2017). Multi-target tracking and data association on road networks using unmanned aerial vehicles. In *2017 IEEE aerospace conference* (pp. 1–11). IEEE.

Bergmann, P., Meinhardt, T., & Leal-Taixe, L. (2019). Tracking without bells and whistles. arXiv preprint arXiv:1903.05625.

Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). Simple online and realtime tracking. In *2016 IEEE international conference on image processing* (ICIP), (pp. 3464–3468). IEEE.

Bewley, A., Ott, L., Ramos, F., & Upcroft, B. (2016). Alextrac: Affinity learning by exploring temporal reinforcement within association chains. In *2016 IEEE international conference on robotics and automation* (ICRA), (pp. 2212–2218). IEEE.

Blackman, S., Busch, M., & Popoli, R. (1995). IMM/MHT tracking and data association for benchmark tracking problem. *4*, In *Proceedings of 1995 american control conference-ACC'95* (pp. 2606–2610). IEEE.

Bozorgtabar, B., & Goecke, R. (2017). MSMCT: Multi-state multi-camera tracker. *IEEE Transactions on Circuits and Systems for Video Technology*, *28*(12), 3361–3376.

Buluttekin, H., Aksu, E., & Bhanuprakash, A. (2017). Comparative study of data association approaches for multiple vehicle tracking. In *2017 21st international conference on system theory, control and computing* (ICSTCC), (pp. 278–284). IEEE.

Cabrera, J. B. (2018). Scheduling variable field-of-view sensors for tracking multiple objects. In *2018 52nd asilomar conference on signals, systems, and computers* (pp. 2174–2178). IEEE.

Cantas, M. R., Chand, A., Zhang, H., Surnilla, G. C., & Guvenc, L. (2021). Data association between perception and V2V communication sensors. arXiv preprint arXiv:2101.08228.

Chalvatzaki, G., Papageorgiou, X. S., Tzafestas, C. S., & Maragos, P. (2018a). Augmented human state estimation using interacting multiple model particle filters with probabilistic data association. *IEEE Robotics and Automation Letters*, *3*(3), 1872–1879.

Chalvatzaki, G., Papageorgiou, X. S., Tzafestas, C. S., & Maragos, P. (2018b). Augmented human state estimation using interacting multiple model particle filters with probabilistic data association. *IEEE Robotics and Automation Letters*, *3*(3), 1872–1879.

Chang, Y.-S., Chiao, H.-T., Abimannan, S., Huang, Y.-P., Tsai, Y.-T., & Lin, K.-M. (2020). An LSTM-based aggregated model for air pollution forecasting. *Atmospheric Pollution Research*, *11*(8), 1451–1463.

Chen, H., Aldea, E., Le Hégarat-Mascle, S., & Despiegel, V. (2020). Use of scene geometry priors for data association in egocentric views. In *2020 8th international workshop on biometrics and forensics* (IWBF), (pp. 1–6). IEEE.

Chen, L., Lou, J., Zhu, W., Xia, Q., & Ren, M. (2017). Multi-appearance segmentation and extended 0-1 program for dense small object tracking. arXiv preprint arXiv:1712.05116.

Chen, T., Pennisi, A., Li, Z., Zhang, Y., & Sahli, H. (2018). A hierarchical association framework for multi-object tracking in airborne videos. *Remote Sensing*, *10*(9), 1347.

Chen, J., Sheng, H., Zhang, Y., & Xiong, Z. (2017). Enhancing detection model for multiple hypothesis tracking. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 18–27).

Chen, Y., Xie, X., Yu, B., Li, Y., & Lin, K. (2021). Multitarget vehicle tracking and motion state estimation using a novel driving environment perception system of intelligent vehicles. *Journal of Advanced Transportation*, *2021*.

Cherian, A., Sra, S., Gould, S., & Hartley, R. (2018). Non-linear temporal subspace representations for activity recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2197–2206).

Choi, H., & Jeon, M. (2016). Data association for non-overlapping multi-camera multi-object tracking based on similarity function. In *2016 IEEE international conference on consumer electronics-asia* (ICCE-Asia), (pp. 1–4). IEEE.

Chong, C.-Y. (2012). Graph approaches for data association. In *2012 15th international conference on information fusion* (pp. 1578–1585). IEEE.

Choutas, V., Weinzaepfel, P., Revaud, J., & Schmid, C. (2018). Potion: Pose motion representation for action recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7024–7033).

Chu, Q., Ouyang, W., Liu, B., Zhu, F., & Yu, N. (2020). Dasot: A unified framework integrating data association and single object tracking for online multi-object tracking. *34*, In *Proceedings of the AAAI conference on artificial intelligence* (07), (pp. 10672–10679).

Dai, P., Wang, X., Zhang, W., & Chen, J. (2018). Instance segmentation enabled hybrid data association and discriminative hashing for online multi-object tracking. *IEEE Transactions on Multimedia*.

Daniłowicz, M. (2020). Multiple object tracking and segmentation with R-CNN networks.

Date, K., Gross, G. A., & Nagi, R. (2014). Test and evaluation of data association algorithms in hard+ soft data fusion. In *17th international conference on information fusion* (FUSION), (pp. 1–8). IEEE.

De Sousa, S., & Kropatsch, W. G. (2015). Graph-based point drift: Graph centrality on the registration of point-sets. *Pattern Recognition*, *48*(2), 368–379.

Dehghan, A., & Shah, M. (2017). Binary quadratic programing for online tracking of hundreds of people in extremely crowded scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(3), 568–581.

Dehghan, A., Tian, Y., Torr, P. H., & Shah, M. (2015). Target identity-aware network flow for online multiple target tracking. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1146–1154).

Dimitrievski, M., Veelaert, P., & Philips, W. (2019). Behavioral pedestrian tracking using a camera and LiDAR sensors on a moving vehicle. *Sensors*, *19*(2), 391.

Doherty, K. J., Baxter, D. P., Schneeweiss, E., & Leonard, J. J. (2020). Probabilistic data association via mixture models for robust semantic SLAM. In *2020 IEEE international conference on robotics and automation* (ICRA), (pp. 1098–1104). IEEE.

Dorai, Y., Chausse, F., Gazzah, S., & Amara, N. E. B. (2017). Multi target tracking by linking tracklets with a convolutional neural network. In *VISIGRAPP (6: VISAPP)* (pp. 492–498).

Duan, C., & Li, X. (2021). Multi-target tracking based on deep sort in traffic scene. *1952*, (2), IOP Publishing, Article 022074,

Elliethy, A., & Sharma, G. (2018). Vehicle tracking in wide area motion imagery via stochastic progressive association across multiple frames. *IEEE Transactions on Image Processing*, *27*(7), 3644–3656.

Emami, P., Pardalos, P. M., Elefteriadou, L., & Ranka, S. (2020). Machine learning methods for data association in multi-object tracking. *ACM Computing Surveys*, *53*(4), 1–34.

Faber, W. R., Hussein, I. I., Kent, J. T., Bhattacharjee, S., & Jah, M. (2018). Optical data processing using directional statistics in a multiple hypothesis framework with maneuvering objects. In 2018 space flight mechanics meeting (pp.1971).

Fagot-Bouquet, L., Audigier, R., Dhome, Y., & Lerasle, F. (2016). Improving multi-frame data association with sparse representations for robust near-online multi-object tracking. In *European conference on computer vision* (pp. 774–790). Springer.

Fan, L., Wang, Z., Cail, B., Tao, C., Zhang, Z., Wang, Y., Li, S., Huang, F., Fu, S., & Zhang, F. (2016). A survey on multiple object tracking algorithm. In *2016 IEEE international conference on information and automation* (ICIA), (pp. 1855–1862). IEEE.

Farazi, H., & Behnke, S. (2017). Online visual robot tracking and identification using deep LSTM networks. In *2017 IEEE/RSJ international conference on intelligent robots and systems* (IROS), (pp. 6118–6125). IEEE.

Farhodov, X., Moon, K.-S., Lee, S.-H., & Kwon, K.-R. (2020). LSTM network with tracking association for multi-object tracking. *Journal of Korea Multimedia Society*, *23*(10), 1236–1249.

Feng, P., Wang, W., Dlay, S., Naqvi, S. M., & Chambers, J. (2016). Social force model-based MCMC-OCSVM particle PHD filter for multiple human tracking. *IEEE Transactions on Multimedia*, *19*(4), 725–739.

Fiaz, M., Mahmood, A., & Jung, S. K. (2018). Tracking noisy targets: A review of recent object tracking approaches. arXiv preprint arXiv:1802.03098.

Gan, W., Wang, S., Lei, X., Lee, M.-S., & Kuo, C.-C. J. (2018). Online CNN-based multiple object tracking with enhanced model updates and identity association. *Signal Processing: Image Communication*, *66*, 95–102.

Gao, S., Han, Z., Li, C., Ye, Q., & Jiao, J. (2015). Real-time multipedestrian tracking in traffic scenes via an RGB-D-based layered graph model. *IEEE Transactions on Intelligent Transportation Systems*, *16*(5), 2814–2825.

Godinez, W. J., & Rohr, K. (2014). Tracking multiple particles in fluorescence time-lapse microscopy images via probabilistic data association. *IEEE Transactions on Medical Imaging*, *34*(2), 415–432.

Gong, Y. (2005). Integrated object detection and tracking by multiple hypothesis analysis. *NEC Journal of Advanced Technology*, *2*(1), 13–18.

Granström, K., Renter, S., Fatemi, M., & Svensson, L. (2017). Pedestrian tracking using Velodyne data—Stochastic optimization for extended object tracking. In *2017 IEEE intelligent vehicles symposium (Iv)* (pp. 39–46). IEEE.

Granström, K., Svensson, L., Reuter, S., Xia, Y., & Fatemi, M. (2017). Likelihood-based data association for extended object tracking using sampling methods. *IEEE Transactions on Intelligent Vehicles*, *3*(1), 30–45.

Gulati, D., Sharif, U., Zhang, F., Clarke, D., & Knoll, A. (2017). Data association—solution or avoidance: Evaluation of a filter based on RFS framework and factor graphs with SME. In *2017 IEEE international conference on multisensor fusion and integration for intelligent systems* (MFI), (pp. 372–377). IEEE.

Guo, Y., & Cheung, N.-M. (2018). Efficient and deep person re-identification using multi-level similarity. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2335–2344).

Guo, S., Wang, J., Wang, X., & Tao, D. (2021). Online multiple object tracking with cross-task synergy. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 8136–8145).

Haag, S., Duraisamy, B., Koch, W., & Dickmann, J. (2018). Classification assisted tracking for autonomous driving domain. In *2018 sensor data fusion: trends, solutions, applications* (SDF), (pp. 1–8). IEEE.

Hamid Rezatofighi, S., Milan, A., Zhang, Z., Shi, Q., Dick, A., & Reid, I. (2015). Joint probabilistic data association revisited. In Proceedings of the IEEE international conference on computer vision (pp. 3047–3055).

He, Y., Han, J., Yu, W., Hong, X., Wei, X., & Gong, Y. (2020). City-scale multi-camera vehicle tracking by semantic attribute parsing and cross-camera tracklet matching. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops (pp. 576–577).

He, J., Huang, Z., Wang, N., & Zhang, Z. (2021). Learnable Graph Matching: Incorporating Graph Partitioning with Deep Feature Learning for Multiple Object Tracking. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 5299–5309).

He, S., Shin, H.-S., & Tsourdos, A. (2020). Trajectory optimization for multitarget tracking using joint probabilistic data association filter. *Journal of Guidance, Control, and Dynamics*, *43*(1), 170–178.

Hou, L., Wan, W., Hwang, J.-N., Muhammad, R., Yang, M., & Han, K. (2017). Human tracking over camera networks: a review. *EURASIP Journal on Advances in Signal Processing*, *2017*(1), 43.

Hou, J., Yang, Y., Wang, Z., & Chen, Y. (2017). Multiple hypothesis tracking in the presence of deception jamming based on multi-feature fusion. In *2017 20th international conference on information fusion* (Fusion), (pp. 1–8). IEEE.

Hu, E., Deng, Z., Jiang, K., & Wu, C. (2020). Kullback-Leibler differential entropy equation based CIMM-PDA for reliable positioning. *Alexandria Engineering Journal*, *59*(4), 2607–2615.

Huang, Y., Chong, S. Y., & Song, T. L. (2017). Track-to-track fusion using multiple detection linear multitarget integrated probabilistic data association. In *ICINCO (1)* (pp. 431–439).

Huang, P., Han, S., Zhao, J., Liu, D., Wang, H., Yu, E., & Kot, A. C. (2020). Refinements in motion and appearance for online multi-object tracking. arXiv preprint arXiv:2003.07177.

Hunde, A., & Ayalew, B. (2018). Automated multi-target tracking in public traffic in the presence of data association uncertainty. In *2018 annual american control conference* (ACC), (pp. 300–306). IEEE.

Hung, W.-C., Kretzschmar, H., Lin, T.-Y., Chai, Y., Yu, R., Yang, M.-H., & Anguelov, D. (2020). SoDA: Multi-object tracking with soft data association. arXiv preprint arXiv:2008.07725.

Jaiswal, A., Beyer, K., Frischknecht, F., & Rohr, K. (2018). Multi-channel boosting and multi-scale localization-based tracking of dense malarial sporozoites. In *2018 IEEE 15th international symposium on biomedical imaging*.

Jeong, J.-M., Yoon, T.-S., & Park, J.-B. (2014). Kalman filter based multiple objects detection-tracking algorithm robust to occlusion. In *2014 proceedings of the sice annual conference* (SICE), (pp. 941–946). IEEE.

Jiang, X., & Cao, X. (2016). Surveillance from above: A detection-and-prediction based multiple target tracking method on aerial videos. In *2016 integrated communications navigation and surveillance* (ICNS), (pp. 4D2–1). IEEE.

Jiang, Z., & Huynh, D. Q. (2017). Multiple pedestrian tracking from monocular videos in an interacting multiple model framework. *IEEE Transactions on Image Processing*, *27*(3), 1361–1375.

Jiang, Z., Huynh, D. Q., Zhang, J., & Wu, Q. (2017). Part-based data association for visual tracking. In *2017 international conference on digital image computing: techniques and applications* (DICTA), (pp. 1–8). IEEE.

Jiang, H., Wang, J., Gong, Y., Rong, N., Chai, Z., & Zheng, N. (2015). Online multi-target tracking with unified handling of complex scenarios. *IEEE Transactions on Image Processing*, *24*(11), 3464–3477.

Joerger, M., & Hassani, A. (2020). A new data association method using Kalman filter innovation vector projections. In *2020 IEEE/ION Position, location and navigation symposium* (PLANS), (pp. 318–327). IEEE.

Kaiser, M., Otte, C., Runkler, T. A., & Ek, C. H. (2018). Data association with Gaussian processes. arXiv preprint arXiv:1810.07158.

Kara, S. F., & Özkan, E. (2018). Multi-ellipsoidal extended target tracking using sequential Monte Carlo. In *2018 21st international conference on information fusion* (FUSION), (pp. 1–8). IEEE.

Kim, C., Fuxin, L., Alotaibi, M., & Rehg, J. M. (2021). Discriminative appearance modeling with multi-track pooling for real-time multi-object tracking. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 9553–9562).

Kokul, T., Ramanan, A., & Pinidiyaarachchi, U. (2015). Online multi-person tracking-by-detection method using ACF and particle filter. In *2015 IEEE seventh international conference on intelligent computing and information systems* (ICICIS), (pp. 529–536). IEEE.

Krishnaswamy, S., & Kumar, M. (2018). A tensor decomposition approach to data association. In 2018 AIAA guidance, navigation, and control conference (pp. 1134).

Kulmon, P., & Stukovska, P. (2018). Assessing multiple-target tracking performance of GNN association algorithm. In *2018 19th international radar symposium* (IRS), (pp. 1–10). IEEE.

Lan, H., Pan, Q., Yang, F., Sun, S., & Li, L. (2016). Variational Bayesian approach for joint multitarget tracking of multiple detection systems. In *2016 19th international conference on information fusion* (FUSION), (pp. 1260–1267). IEEE.

Lázaro-Gredilla, M., Van Vaerenbergh, S., & Lawrence, N. D. (2012). Overlapping mixtures of Gaussian processes for the data association problem. *Pattern Recognition*, *45*(4), 1386–1395.

Leal-Taixé, L., Canton-Ferrer, C., & Schindler, K. (2016). Learning by tracking: Siamese CNN for robust target association. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 33–40).

Lee, K.-H., & Hwang, J.-N. (2015). On-road pedestrian tracking across multiple driving recorders. *IEEE Transactions on Multimedia*, *17*(9), 1429–1438.

Lee, K.-H., Kanzawa, Y., Derry, M., & James, M. R. (2018). Multi-target track-to-track fusion based on permutation matrix track association. In *2018 IEEE intelligent vehicles symposium (IV)* (pp. 465–470). IEEE.

Lee, S.-H., Kim, M.-Y., & Bae, S.-H. (2018). Learning discriminative appearance models for online multi-object tracking with appearance discriminability measures. *IEEE Access*, *6*, 67316–67328.

Lee, G., Mallipeddi, R., & Lee, M. (2017). Trajectory-based vehicle tracking at low frame rates. *Expert Systems with Applications*, *80*, 46–57, URL http://dx.doi.org/10.1016/j.eswa.2017.03.023.

Lee, E., Zhang, Q., & Song, T. (2017). Markov chain realization of joint integrated probabilistic data association. *Sensors*, *17*(12), 2865.

Leonard, M. R., & Zoubir, A. M. (2019). Multi-target tracking in distributed sensor networks using particle PHD filters. *Signal Processing*, *159*, 130–146.

Li, L.-Q., Li, E.-Q., & He, W.-M. (2017). A novel fuzzy data association approach for visual multi-object tracking. In *ITM web of conferences, Vol. 12* (p. 05004). EDP Sciences.

Li, H., Liu, Y., Lin, W., Xu, L., & Wang, J. (2020). Data association methods via video signal processing in imperfect tracking scenarios: A review and evaluation. *Mathematical Problems in Engineering*, *2020*.

Li, M., Liu, Z., Xiong, Y., & Li, Z. (2017). Multi-person tracking by discriminative affinity model and hierarchical association. In *2017 3rd IEEE International Conference on Computer and Communications* (ICCC), (pp. 1741–1745). IEEE.

Li, P., Wang, D., Wang, L., & Lu, H. (2018). Deep visual tracking: Review and experimental comparison. *Pattern Recognition*, *76*, 323–338.

Liang, T., Lan, L., & Luo, Z. (2020). Enhancing the association in multi-object tracking via neighbor graph. arXiv preprint arXiv:2007.00265.

Lin, C.-C., & Hung, Y. (2018). A prior-less method for multi-face tracking in unconstrained videos. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 538–547).

Lipovits, A., Czúni, L., Tömördi, K., & Vörösházi, Z. (2021). *Multiple object tracking by bounding boxes without using texture information and optical flow*. Václav Skala-UNION Agency.

Liu, J., Cao, X., Li, Y., & Zhang, B. (2017). Online multi-object tracking using hierarchical constraints for complex scenarios. *IEEE Transactions on Intelligent Transportation Systems*, *19*(1), 151–161.

Liu, Q., Chu, Q., Liu, B., & Yu, N. (2020). GSM: Graph similarity model for multi-object tracking. In *IJCAI* (pp. 530–536).

Liu, M., Jin, C.-B., Yang, B., Cui, X., & Kim, H. (2018). Online multiple object tracking using confidence score-based appearance model learning and hierarchical data association. *IET Computer Vision*, *13*(3), 312–318.

Liu, W., Liu, Y., Gunawan, B. A., & Bucknall, R. (2020). Practical moving target detection in maritime environments using fuzzy multi-sensor data fusion. *International Journal of Fuzzy Systems*, 1–19.
Liu, Y., Yao, L., Xiong, W., & Zhou, Z. (2018). Joint kinematic and feature tracking of ships with satellite electronic information. *The Journal of Navigation*, *71*(5), 1178–1194.
Liu, H., Zhang, H., & Mertz, C. (2019). DeepDA: LSTM-based deep data association network for multi-targets tracking in clutter. arXiv preprint arXiv:1907.09915.
Luo, Z., Attari, M., Habibi, S., & Von Mohrenschildt, M. (2019). Online multiple maneuvering vehicle tracking system based on multi-model smooth variable structure filter. *IEEE Transactions on Intelligent Transportation Systems*.
Luo, W., Xing, J., Milan, A., Zhang, X., Liu, W., Zhao, X., & Kim, T.-K. (2014). Multiple object tracking: A literature review. arXiv preprint arXiv:1409.7618.
Luvizon, D. C., Picard, D., & Tabia, H. (2018). 2d/3d pose estimation and action recognition using multitask deep learning. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 5137–5146).
Mahemuti, B., Inoue, D., Kakugo, A., & Konagaya, A. (2016). Investigation of the microtubule dynamics with probabilistic data association filter. In *2016 IEEE 11th annual international conference on nano/micro engineered and molecular systems* (NEMS), (pp. 101–106). IEEE.
Mandal, V., & Adu-Gyamfi, Y. (2020). Object detection and tracking algorithms for vehicle counting: a comparative analysis. *Journal of Big Data Analytics in Transportation*, *2*(3), 251–261.
Mei, W., Xiong, G., Gong, J., Yong, Z., Chen, H., & Di, H. (2017). Multiple moving target tracking with hypothesis trajectory model for autonomous vehicles. In *2017 IEEE 20th international conference on intelligent transportation systems* (ITSC), (pp. 1–6). IEEE.
Meinhardt, T., Kirillov, A., Leal-Taixe, L., & Feichtenhofer, C. (2021). Trackformer: Multi-object tracking with transformers. arXiv preprint arXiv:2101.02702.
Meneses, M., Matos, L., Prado, B., de Carvalho, A., & Macedo, H. (2020). Learning to associate detections for real-time multiple object tracking. arXiv preprint arXiv:2007.06041.
Meng, J., Wu, A., & Zheng, W.-S. (2019). Deep asymmetric video-based person re-identification. *Pattern Recognition*, *93*, 430–441.
Michaelis, M., Berthold, P., Meissner, D., & Wuensche, H.-J. (2017). Heterogeneous multi-sensor fusion for extended objects in automotive scenarios using Gaussian processes and a GMPHD-filter. In *2017 sensor data fusion: trends, solutions, applications* (SDF), (pp. 1–6). IEEE.
Milan, A., Roth, S., & Schindler, K. (2013). Continuous energy minimization for multitarget tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *36*(1), 58–72.
Niedfeldt, P. C., Ingersoll, K., & Beard, R. W. (2017). Comparison and analysis of recursive-RANSAC for multiple target tracking. *IEEE Transactions on Aerospace and Electronic Systems*, *53*(1), 461–476.
Noh, G.-S., & Jeon, M. (2015). A systematic framework for real-time online multi-object tracking. In *2015 international conference on control, automation and information sciences* (ICCAIS), (pp. 57–61). IEEE.
Ooi, H.-L., Bilodeau, G.-A., & Saunier, N. (2020). Supervised and unsupervised detections for multiple object tracking in traffic scenes: A comparative study. In *International conference on image analysis and recognition* (pp. 42–55). Springer.
Pang, J., Qiu, L., Li, X., Chen, H., Li, Q., Darrell, T., & Yu, F. (2021). Quasi-dense similarity learning for multiple object tracking. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 164–173).
Park, C., Woehl, T. J., Evans, J. E., & Browning, N. D. (2014). Minimum cost multi-way data association for optimizing multitarget tracking of interacting objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *37*(3), 611–624.
Piao, S., Sutjaritvorakul, T., & Berns, K. (2016). Compact data association in multiple object tracking: pedestrian tracking on mobile vehicle as case study. *IFAC-PapersOnLine*, *49*(15), 175–180.
Punchihewa, Y. G., Vo, B.-T., Vo, B.-N., & Kim, D. Y. (2018). Multiple object tracking in unknown backgrounds with labeled random finite sets. *IEEE Transactions on Signal Processing*, *66*(11), 3040–3055.
Qu, C., Zhang, Y., Zhang, X., & Yang, Y. (2020). Reinforcement learning-based data association for multiple target tracking in clutter. *Sensors*, *20*(22), 6595.
Raboaca, M. S., Dumitrescu, C., & Manta, I. (2020). Aircraft trajectory tracking using radar equipment with fuzzy logic algorithm. *Mathematics*, *8*(2), 207.
Rangesh, A., Maheshwari, P., Gebre, M., Mhatre, S., Ramezani, V., & Trivedi, M. M. (2021). TrackMPNN: A message passing graph neural architecture for multi-object tracking. arXiv preprint arXiv:2101.04206.
Rasmussen, C., & Hager, G. D. (2001). Probabilistic data association methods for tracking complex visual objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6), 560–576.
Riahi, D., & Bilodeau, G.-A. (2015). Multiple object tracking based on sparse generative appearance modeling. In *2015 IEEE international conference on image processing* (ICIP), (pp. 4017–4021). IEEE.
Ribeiro, M. I. (2004). Kalman and extended kalman filters: Concept, derivation and properties. *Institute for Systems and Robotics*, *43*.
Ritter, C., Imle, A., Lee, J. Y., Müller, B., Fackler, O. T., Bartenschlager, R., & Rohr, K. (2018). Two-filter probabilistic data association for tracking of virus particles in fluorescence microscopy images. In *2018 IEEE 15th international symposium on biomedical imaging* (ISBI 2018), (pp. 957–960). IEEE.
Sahbani, B., & Adiprawita, W. (2016). Kalman filter and iterative-hungarian algorithm implementation for low complexity point tracking as part of fast multiple object tracking system. In *2016 6th international conference on system engineering and technology* (ICSET), (pp. 109–115). IEEE.
Salscheider, N. O. (2021). Object tracking by detection with visual and motion cues. arXiv preprint arXiv:2101.07549.
Salvi, D., Waggoner, J., Temlyakov, A., & Wang, S. (2013). A graph-based algorithm for multi-target tracking with occlusion. In *2013 IEEE workshop on applications of computer vision* (WACV), (pp. 489–496). IEEE.
Seong, Y.-M., & Park, H. (2012). Multiple target tracking using cognitive data association of spatiotemporal prediction and visual similarity. *Pattern Recognition*, *45*(9), 3451–3462.
Shen, Y., Li, H., Xiao, T., Yi, S., Chen, D., & Wang, X. (2018). Deep group-shuffling random walk for person re-identification. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2265–2274).
Sheng, H., Chen, J., Zhang, Y., Ke, W., Xiong, Z., & Yu, J. (2018). Iterative multiple hypothesis tracking with tracklet-level association. *IEEE Transactions on Circuits and Systems for Video Technology*.
Singh, G., Rajan, S., & Majumdar, S. (2017). A greedy data association technique for multiple object tracking. In *2017 IEEE third international conference on multimedia big data* (BigMM), (pp. 177–184). IEEE.
Stauch, J., Bessell, T., Rutten, M., Baldwin, J., Jah, M., & Hill, K. (2017). Joint probabilistic data association and smoothing applied to multiple space object tracking. *Journal of Guidance, Control, and Dynamics*, *41*(1), 19–33.
Steyer, S., Tanzmeister, G., Lenk, C., Dallabetta, V., & Wollherr, D. (2018). Data association for grid-based object tracking using particle labeling. In *2018 21st international conference on intelligent transportation systems* (ITSC), (pp. 3036–3043). IEEE.
Sun, S., Akhtar, N., Song, H., Mian, A. S., & Shah, M. (2019). Deep affinity network for multiple object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
Sun, P., Jiang, Y., Zhang, R., Xie, E., Cao, J., Hu, X., Kong, T., Yuan, Z., Wang, C., & Luo, P. (2020). Transtrack: Multiple-object tracking with transformer. arXiv preprint arXiv:2012.15460.
Taalimi, A., & Qi, H. (2015). Robust multi-object tracking using confident detections and safe tracklets. In *2015 IEEE international conference on image processing* (ICIP), (pp. 1638–1642). IEEE.
Tafti, A. D., & Sadati, N. (2010). Modified maximum entropy fuzzy data association filter. *Journal of Dynamic Systems, Measurement, and Control*, *132*(2), Article 021013.
Tan, L., Dong, X., Ma, Y., & Yu, C. (2018). A multiple object tracking algorithm based on YOLO detection. In *2018 11th international congress on image and signal processing, biomedical engineering and informatics* (CISP-BMEI), (pp. 1–5). IEEE.
Tang, Z., Gu, R., & Hwang, J.-N. (2018). Joint multi-view people tracking and pose estimation for 3D scene reconstruction. In *2018 IEEE international conference on multimedia and expo* (ICME), (pp. 1–6). IEEE.
Tang, Z., & Hwang, J.-N. (2019). Moana: An online learned adaptive appearance model for robust multiple object tracking in 3d. *IEEE Access*, *7*, 31934–31945.
Tang, Z., Wang, G., Xiao, H., Zheng, A., & Hwang, J.-N. (2018). Single-camera and inter-camera vehicle tracking and 3D speed estimation based on fusion of visual and semantic features. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 108–115).
Tian, Y., Dehghan, A., & Shah, M. (2018). On detection, data association and segmentation for multi-target tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
Tran, A. T., & Harada, K. (2013). Depth-aided tracking multiple objects under occlusion. *Journal of Signal and Information Processing*, *4*(03), 299.
Vivone, G., Braca, P., & Errasti-Alcala, B. (2015). Extended target tracking applied to X-band marine radar data. In *OCEANS 2015-Genova* (pp. 1–6). IEEE.
Wan, X., Wang, J., Kong, Z., Zhao, Q., & Deng, S. (2018). Multi-object tracking using online metric learning with long short-term memory. In *2018 25th IEEE international conference on image processing* (ICIP), (pp. 788–792). IEEE.
Wang, J., Ancha, S., Chen, Y.-T., & Held, D. (2020). Uncertainty-aware self-supervised 3D data association. In *2020 IEEE/RSJ international conference on intelligent robots and systems* (IROS), (pp. 8125–8132). IEEE.
Wang, J., Guo, Y., Tang, X., Hu, Q., & An, W. (2018). Semi-online multiple object tracking using graphical tracklet association. *IEEE Signal Processing Letters*, *25*(11), 1725–1729.
Wang, B., Wang, G., Chan, K. L., & Wang, L. (2016). Tracklet association by online target-specific metric learning and coherent dynamics estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(3), 589–602.
Wang, B. H., Wang, Y., Weinberger, K. Q., & Campbell, M. (2018). Deep person re-identification for probabilistic data association in multiple pedestrian tracking. arXiv preprint arXiv:1810.08565.
Wang, H., Wang, X., Zheng, J., Deller, J. R., Peng, H., Zhu, L., Chen, W., Li, X., Liu, R., & Bao, H. (2014). Video object matching across multiple non-overlapping camera views based on multi-feature fusion and incremental learning. *Pattern Recognition*, *47*(12), 3841–3851.
Wang, C., Xie, X., & Liao, C. (2021). An adaptive fusion model based on Kalman filtering and LSTM for fast tracking of road signs. In *2020 25th international conference on pattern recognition* (ICPR), (pp. 1414–1421). IEEE.

Wang, Z., Zheng, L., Liu, Y., Li, Y., & Wang, S. (2020). Towards real-time multi-object tracking. In *Computer vision–ECCV 2020: 16th european conference*. Glasgow, UK.
Weng, X., Wang, Y., Man, Y., & Kitani, K. (2020). Graph neural networks for 3D multi-object tracking. arXiv preprint arXiv:2008.09506.
Wong, L. L., Kaelbling, L. P., & Lozano-Pérez, T. (2015). Data association for semantic world modeling from partial views. *International Journal of Robotics Research*, *34*(7), 1064–1082.
Wu, S., & Hong, L. (2005). Hand tracking in a natural conversational environment by the interacting multiple model and probabilistic data association (IMM-PDA) algorithm. *Pattern Recognition*, *38*(11), 2143–2158.
Wu, H., Hu, Y., Wang, K., Li, H., Nie, L., & Cheng, H. (2019). Instance-aware representation learning and association for online multi-person tracking. *Pattern Recognition*, *94*, 25–34.
Wu, H., & Li, W. (2016). Robust online multi-object tracking based on KCF trackers and reassignment. In *2016 IEEE global conference on signal and information processing* (GlobalSIP), (pp. 124–128). IEEE.
Xi-yang, Z., Xiao-li, W., & Liang-qun, L. (2018). Online multi-object tracking via maximum entropy intuitionistic fuzzy data association. In *2018 14th IEEE international conference on signal processing* (ICSP), (pp. 803–806). IEEE.
Xiang, J., Zhang, G., & Hou, J. (2019). Online multi-object tracking based on feature representation and Bayesian filtering within a deep learning architecture. *IEEE Access*, *7*, 27923–27935.
Xiao, P., & Zhong, L. (2017). Tracking of non-dividing cells by using generalized voronoi diagram. In *2017 39th annual international conference of the ieee engineering in medicine and biology society* (EMBC), (pp. 2684–2687). IEEE.
Xu, C., & Zhou, Y. (2018). Hierarchical online multi-person pose tracking with multiple cues. In *International conference on neural information processing* (pp. 318–328). Springer.
Yang, F., Chang, X., Dang, C., Zheng, Z., Sakti, S., Nakamura, S., & Wu, Y. (2020). ReMOTS: Self-supervised refining multi-object tracking and segmentation. arXiv preprint arXiv:2007.03200.
Yang, F., Soroush, F., Deng, G., Yu, S., Chu, P., Kiani, M. F., & Ling, H. (2018). Multiple neutrophils tracking in vitro array using high-order temporal information. In *2018 40th annual international conference of the ieee engineering in medicine and biology society* (EMBC), (pp. 1–4). IEEE.
Yang, S., Thormann, K., & Baum, M. (2018). Linear-time joint probabilistic data association for multiple extended object tracking. In *2018 IEEE 10th sensor array and multichannel signal processing workshop* (SAM), (pp. 6–10). IEEE.
Yang, M., Wu, Y., & Jia, Y. (2017). A hybrid data association framework for robust online multi-object tracking. *IEEE Transactions on Image Processing*, *26*(12), 5667–5679.
Yao, Y., Smal, I., & Meijering, E. (2018). Deep neural networks for data association in particle tracking. In *2018 IEEE 15th international symposium on biomedical imaging* (ISBI 2018), (pp. 458–461). IEEE.
Yarkony, J., Adulyasak, Y., Singh, M., & Desaulniers, G. (2020). Data association via set packing for computer vision applications. *Informs Journal on Optimization*, *2*(3), 167–191.
Yingyi, L., Xin, L., Zhenyu, H., & Xinge, Y. (2017). Multiple object tracking by incorporating a particle filter into the min-cost flow model. In *2017 international conference on security, pattern analysis, and cybernetics* (SPAC), (pp. 106–111). IEEE.
Yoon, J. H., Lee, C.-R., Yang, M.-H., & Yoon, K.-J. (2019). Structural constraint data association for online multi-object tracking. *International Journal of Computer Vision*, *127*(1), 1–21.
Yoon, K., Song, Y.-m., & Jeon, M. (2018). Multiple hypothesis tracking algorithm for multi-target multi-camera tracking with disjoint views. *IET Image Processing*, *12*(7), 1175–1184.
Yoon, J. H., Yang, M.-H., Lim, J., & Yoon, K.-J. (2015). BayesIan multi-object tracking using motion context from multiple objects. In *2015 IEEE winter conference on applications of computer vision* (pp. 33–40). IEEE.
Yu, H., Li, G., Su, L., Zhong, B., Yao, H., & Huang, Q. (2020). Conditional GAN based individual and global motion fusion for multiple object tracking in UAV videos. *Pattern Recognition Letters*, *131*, 219–226.
Yuan, T., Krishnan, K., Chen, Q., Breu, J., Roth, T. B., Duraisamy, B., Weiss, C., Maile, M., & Gern, A. (2017). Object matching for inter-vehicle communication systems—An IMM-based track association approach with sequential multiple hypothesis test. *IEEE Transactions on Intelligent Transportation Systems*, *18*(12), 3501–3512.
Yuan, T., Krishnan, K., Duraisamy, B., Maile, M., & Schwarz, T. (2017). Extended object tracking using IMM approach for a real-world vehicle sensor fusion system. In *2017 IEEE international conference on multisensor fusion and integration for intelligent systems* (MFI), (pp. 638–643). IEEE.
Zeng, Q., Wen, G., & Li, D. (2016). Multi-target tracking by detection. In *2016 international conference on audio, language and image processing* (ICALIP), (pp. 370–374). IEEE.
Zeng, D., Xiong, L., Yu, Z., Chen, Q., Fu, Z., Li, Z., Zhang, P., Xu, P., Qian, Z., & Xiao, H. (2020). A priority data association policy for multitarget tracking on intelligent vehicle risk assessment. *Remote Sensing*, *12*(19), 3255.
Zhang, C., Huang, Y., Wang, Z., Jiang, H., & Yan, D. (2018). Cross-camera multi-person tracking by leveraging fast graph mining algorithm. *Journal of Visual Communication and Image Representation*, *55*, 711–719.
Zhang, J., Li, W., Ogunbona, P. O., Wang, P., & Tang, C. (2016). RGB-D-based action recognition datasets: A survey. *Pattern Recognition*, *60*, 86–105.
Zhang, Y., Liu, M., Liu, X., & Wu, T. (2020). A group target tracking algorithm based on topology. *1544*, In *Journal of Physics: Conference Series*. (1), IOP Publishing, Article 012025.
Zhang, H., Liu, H.-j., & Wang, C.-l. (2019). Learning to multi-target tracking in dense clutter environment with JPDA-recurrent neural networks. *1207*, In *Journal of Physics: Conference Series*. (1), IOP Publishing, Article 012011.
Zhang, T.-h., & Tang, C.-W. (2018). Multiple-target tracking on mixed images with reflections and occlusions. *Journal of Visual Communication and Image Representation*, *52*, 45–57.
Zhong, X., Tay, W. P., Leng, M., Razul, S. G., & See, C. M. S. (2016). Tdoa-fdoa based multiple target detection and tracking in the presence of measurement errors and biases. In *2016 IEEE 17th international workshop on signal processing advances in wireless communications* (SPAWC), (pp. 1–6). IEEE.
Zhu, H., Yuen, K.-V., Mihaylova, L., & Leung, H. (2017). Overview of environment perception for intelligent vehicles. *IEEE Transactions on Intelligent Transportation Systems*, *18*(10), 2584–2601.

**Lionel Rakai** received his B.S. as a double major in Computer Science and Information Systems and M.S degree in Information Systems from The University of the South Pacific and is currently working towards a Ph.D. degree in intelligent transportation and information system engineering with Chang'an University. His research interests include deep learning, multiple object tracking, data association, and action recognition.

**Huansheng Song** received the B.S. and M.S. degrees in communication and electronic systems and the Ph.D. degree in information and communication engineering from Xian Jiaotong University, Xian, China, in 1985, 1988, and 1996, respectively. Since 2004, he has been with the Information Engineering Institute, Changan University, Xian, where he became a Professor in 2006 and was nominated as the Dean in 2012. He has been involved in research on intelligent transportation systems for many years and has led a research team to develop a vehicle licence plate reader and a traffic event detection system based on videos, which has brought about complete industrialization. His current research interests include image processing and recognition, as well as intelligent transportation systems.

**ShiJie Sun** received his Ph.D in intelligent transportation and information system engineering from the University of Chang'an University, China, in 2019. He has been a visiting (joint) Ph.D. candidate at the University of Western Australia in 2017. He is a reviewer of the top journals an conferences such as TPAMI, CVPR and ECCV. His research interests include machine learning, object detection, localization & tracking, and action recognition.

**Wentao Zhang** received his B.S degree in electronic information engineering from Shangluo university, and received the M.S degree in information and communication engineering from Chang'an University. He is currently working towards the Ph.D. degree in transportation information engineering and control with Chang'an University. His research interests include deep learning, object detection& tracking, action recognition.

**Yanni Yang** received the B.S. degree in Measurement and Control Technology and Instrumentation from Hebei University of Engineering, and the M.S. degree in Control Science and Engineering from Xinjiang University. She is currently working towards the Ph.D. degree in Intelligent Transportation and Information System Engineering with Chang'an University. Her research interests include computer vision, deep learning, and object detection.